



# Information, institutions et efficacité: essais en économie expérimentale

Adam Zylbersztein

## ► To cite this version:

Adam Zylbersztein. Information, institutions et efficacité: essais en économie expérimentale. Economies et finances. Université Panthéon-Sorbonne - Paris I, 2013. Français. NNT: 2013PA010015 . tel-00984244

**HAL Id: tel-00984244**

**<https://theses.hal.science/tel-00984244>**

Submitted on 28 Apr 2014

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITÉ PARIS 1 - PANTHÉON SORBONNE  
U.F.R DE SCIENCES ECONOMIQUES

Année 2012-2013

Numéro attribué par la bibliothèque

| | | | | | | | | | | |

**Thèse pour le doctorat de Sciences Economiques**

*soutenue publiquement par*

**ADAM ZYLBERSZTEJN**

*le 11 juin 2013*

INFORMATION, INSTITUTIONS AND EFFICIENCY:  
AN EXPERIMENTAL APPROACH

***Directeurs de thèse***

Nicolas Jacquemet  
Jean-Marc Tallon

Professeur à l'Université de Lorraine  
Directeur de Recherche au CNRS

***Jury***

Juergen Bracht  
Frédéric Koessler  
Antoine Terracol  
Marie-Claire Villeval

Professeur, Université d'Aberdeen  
Directeur de Recherche, CNRS  
Maître de Conférences, Université Paris 1 Panthéon-Sorbonne  
Directrice de Recherche, CNRS

# Contents

<b>1</b>	<b>Inequality aversion, information, and coordination</b>	<b>21</b>
1.1	Introduction . . . . .	21
1.2	Inequality aversion and efficiency in coordination games . . . . .	23
1.2.1	Empirical strategy . . . . .	24
1.2.2	Statistical procedure for mean comparisons . . . . .	28
1.2.3	Results . . . . .	30
1.3	Coordination under enhanced information . . . . .	38
1.3.1	Analysis of the game . . . . .	38
1.3.2	Experimental design . . . . .	41
1.3.3	Results . . . . .	44
1.4	Summary and conclusion . . . . .	57
1.5	Supplementary material . . . . .	58
<b>2</b>	<b>Coordination with communication under oath</b>	<b>70</b>
2.1	Introduction . . . . .	70
2.2	Description of the experiment . . . . .	72
2.2.1	Experimental design . . . . .	74
2.2.2	Experimental procedures . . . . .	76
2.3	Results . . . . .	77
2.3.1	Communication under oath . . . . .	78
2.4	Commitment without communication . . . . .	81
2.4.1	Experimental design . . . . .	82
2.4.2	Results . . . . .	83
2.4.3	Communication under oath and strategic uncertainty . . . . .	84
2.5	Conclusion . . . . .	86
2.5.1	Bounded rationality and strategic uncertainty . . . . .	88
<b>3</b>	<b>Ex post communication in a public goods game</b>	<b>91</b>
3.1	Introduction . . . . .	91

3.2	Empirical strategy . . . . .	93
3.2.1	Experimental game and conditions . . . . .	95
3.2.2	Experimental procedures. . . . .	96
3.3	Results . . . . .	97
3.3.1	Behavior with and without ex post communication . . . . .	97
3.3.2	Formation of ex-post messages . . . . .	103
3.4	Summary and discussion . . . . .	103

# List of Figures

1.1	Baseline game . . . . .	65
1.2	Communication . . . . .	65
1.3	Observation . . . . .	66
1.4	Evolution of outcomes over rounds . . . . .	67
2.1	Oath form used in the experiment . . . . .	75
2.2	Decisions of receivers and senders by treatment . . . . .	79
2.3	Communication behavior of senders by treatment . . . . .	80
2.4	Truthfulness of senders by treatment (Empirical distribution function) . . . . .	81
2.5	Messages trusted by player As (Empirical distribution function) . . . . .	82
2.6	Observed decisions without communication . . . . .	84
2.7	Efficient coordination in human subjects and automated players treatments . . . . .	85
2.8	Player As' behavior against automated players . . . . .	87
2.9	Proportion of $R$ decisions across rounds and treatments . . . . .	89
3.1	Distribution of contributions across treatments . . . . .	99
3.2	Contributions, ex-post communication and re-matching . . . . .	100
3.3	Sent disapproval points and the relative size of contribution . . . . .	102
3.3	Evolution of average contribution by game and treatment . . . . .	111

# List of Tables

1	Rosenthal's coordination game . . . . .	15
1.1	Summary of experimental evidence on Rosenthal's game . . . . .	23
1.2	Generic game . . . . .	24
1.3	Overview of experimental treatments . . . . .	26
1.4	Observed behavior in Treatments 1 . . . . .	31
1.5	Observed decisions . . . . .	32
1.6	Distribution of players according to their actions . . . . .	34
1.7	Parametric regressions on the determinants of cooperative behavior . . . . .	35
1.8	The experimental game . . . . .	38
1.9	Summary of experimental evidence on related games . . . . .	44
1.10	Overall effects of information treatments . . . . .	45
1.11	Statistical support to Table 1.10 . . . . .	47
1.12	Informational content of signals . . . . .	49
1.13	Statistical support to Table 1.12 . . . . .	50
1.14	Marginal effect of components of past history in the observation treatment . . . . .	52
1.15	Probit regressions on player As' reliance in the three experimental games . . . . .	53
1.16	Average payoffs from the use of dominant strategies . . . . .	54
1.17	Baseline game: evolution of frequencies of choices and outcomes over time . . . . .	68
1.18	Communication game: evolution of frequencies of choices and outcomes over time . . . . .	68
1.19	Observation game: evolution of frequencies of choices and outcomes over time . . . . .	69
2.1	The experimental game . . . . .	72
2.2	Overview of existing experimental evidence . . . . .	73
2.3	Aggregate results . . . . .	77
2.4	Coordination under oath, without communication: Aggregate results . . . . .	83
2.5	The experimental games . . . . .	88
3.1	Average contributions according to treatment and experimental game . . . . .	98
3.2	Determinants of individual and group behavior . . . . .	101

3.3	Sent disapproval points and the relative size of contribution . . . . .	104
3.4	Your gain in ECU in a given period as a function of your decision and your group member's decision . . . . .	108
3.5	Determinants of individual and group behavior: marginal effects . . . . .	112

# Acknowledgments

Leonardo da Vinci's *Mona Lisa* has been created by few and it is known to many. A thesis in economics is the exact opposite of *Mona Lisa*. This thesis is no exception, and many people should be credited for supporting my work during the last four years.

Nicolas Jacquemet and Jean-Marc Tallon, my PhD advisors, devoted a substantial amount of time and effort for putting my scientific endeavors on the right track. A long and fruitful collaboration with Nicolas made me the experimental economist I am today, taught how to ask questions and look for answers. His usual prelude to our discussions – "imagine that we live in a world where ..." – has become my own over time and, I am sure, will remain handy in actually understanding the world we live in. Jean-Marc's emphasis on seeking a broad picture in experimental results has affected the way I think about experimental economics and economic research as such.

I would also like to acknowledge all the help I have received from my thesis committee – Juergen Bracht, Frédéric Koessler, Antoine Terracol and Marie-Claire Villeval – who have shown the patience and the courage to read my papers, and have provided me with great feedback throughout the intense process of writing up of this thesis.

Credits go to my co-authors – Nobuyuki Hanaki, Stéphane Luchini and Jason Shogren (and Nicolas, of course) – for our inspiring collaboration, and your huge contribution to the research presented in this thesis.

These four years would not be the same without a bunch of PhD students from the fourth floor. In particular, many thanks to Batool, Nicolas, Noémi and Sébastien for bearing with me for such a long period of time in our cosy office 402, and to Hela for all the gossip-filled chats we used to have.

One of the most inspiring experiences in my life, and a turning point for this thesis, was the visit at Columbia University. I would like to thank Pierre-André Chiappori for having kindly agreed to host me, and for all his assistance during my stay. This research visit would not have been possible without financial and administrative support from the Alliance Program and the Collège des Ecoles Doctorales de l'Université Paris 1 Panthéon-Sorbonne, which I gratefully acknowledge. The time spent in The Big Apple would never be the same without Uncle Marty – many thanks for your fantastic hospitality, I will never forget about it!



Last but not least, many thanks to my family – Alexandra and Vincent in Paris, and Michal, Maria and Anna in Warsaw, for their continuous support which helped me cope with all ups and downs that came along the way during these four years. A special recognition goes to Vincent, who pointed out a glaring shortcoming of this thesis: I was writing it for so long and finally forgot to include an airplane. Thanks for the tip, here is the plane:



# Introduction

One of the traditional research topics in public economics is the question of superseding "the tragedy of the commons" (Hardin, 1968) by "the governance of commons" (Ostrom, 1990). As an insightful illustration, take the famous case study on the self-governance of a common pool resource by Elinor Ostrom. As described by Ostrom, Turkish fishermen from the region of Antalya share fisheries that contain two types of fishing grounds: those that are rich in fish and those that are scarce of fish. Although a clear preference of *each* vessel is to go to a rich fishing spot rather than a poor one, such behavior among *all* vessels would eventually kill the resource. As a solution to this issue, these fishermen have implemented a rotational mechanism that attributes to each vessel a certain amount of time in both kinds of fisheries. This solution *(i)* allows all vessels to benefit equally from the common resource, while *(ii)* enabling the shoals to rebuild – which sustains the resource over time.

This example is very helpful for describing the focus of the present thesis. First, the case studied by Ostrom involves two important types of strategic interactions in which institutions may help reduce inefficiencies. On the one hand, these economic agents face a *cooperation problem*: they must agree to conform their personal motives (*i.e.*, always go to the rich fisheries) with the common interest (*i.e.*, select fishing spots so as to preserve the resource). On the other hand, they also need to overcome an important *coordination problem* by establishing a way to synchronize their actions. Second, the institutional design they adopt can be characterized as non-monetary in the sense that it involves no pecuniary incentives (such as awards and fines), but is rather based on the availability of information on the nature of fisheries and on how others exploit them, as well as on the personal commitment of all participants.

In the present thesis, we aim at understanding when and why such institutions – based on the transmission of information between economic agents and their personal commitment – contribute to enhancing the efficiency of human interactions. In line with recent methodological advances in economic research, our empirical strategy relies on laboratory experiments which enable us to acquire causal knowledge by means of controlled variation (Falk and Heckman, 2009). More precisely, we *(i)* construct decision-making tasks that exhibit relevant strategic features and display desirable game theoretical properties, and *(ii)* collect empirical data in the lab – where human subjects face these tasks in a carefully crafted decision-making environment.

## Motivation of the present work

A fundamental concept used to construct theoretical predictions of human interactions is the Nash equilibrium: a state in which every player makes the best decision given all the other players' decisions. However, a notable shortfall of this concept is that the (pure strategy) Nash equilibrium is not necessary unique. Moreover, in certain strategic contexts the non-equilibrium outcomes should not be considered irrelevant: for instance, when the non-Nash outcome is Pareto superior with respect to the equilibrium outcome (Luce and Raiffa, 1957; Harrison and Hirshleifer, 1989). The interplay of the two solution principles – the Nash equilibrium and the Pareto efficiency – is especially pronounced in two types of economic games: coordination games and the public goods games.

### Coordination games: multiplicity of equilibria and strategic uncertainty

Let us first introduce the notion of a *coordination game* which constitutes the basis of the analysis conducted in Chapters 1 and 2 of the present thesis. In simple words, a coordination game illustrates an interactive situation with strategic complementarity: agents may create additional value simply by synchronizing their individual actions. These games have been found useful for modeling many problems in macroeconomics and industrial organization (Cooper, DeJong, Forsythe, and Ross, 1990). In the broadest sense, a coordination game is a game involving multiple Nash equilibria (Schelling, 1960). Some authors prefer to narrow this definition and additionally posit the Pareto rankability of the Nash equilibria (Cooper, DeJong, Forsythe, and Ross, 1990); we will henceforth line with this approach.

In coordination games where players do not have dominant strategies (such as the stag hunt game, the minimum-effort game, and other order-statistic games), decisions are being made under uncertainty about others' intentions, and thus one is usually interested in establishing a *focal point* – an equilibrium which players will naturally focus on. Clearly, the most appealing focal point is the Pareto dominant (or the *payoff dominant*) Nash equilibrium (Schelling, 1960). However, some game theorists claim that the riskiness of equilibrium is also an important factor, and evoke the least risky (or the *risk dominant*) equilibrium as an alternative focal point (Harsanyi and Selten, 1988). Experimental data for this class of games suggest that (i) the Nash equilibrium is a reliable predictor of the actual outcomes (Cooper, DeJong, Forsythe, and Ross, 1990), and that (ii) the risk dominant equilibrium happens to be at least as a meaningful focal point as the payoff dominant equilibrium (Camerer, 2003; Devetag and Ortmann, 2007). On the other hand, coordination games may also exhibit (weakly) dominant strategies. Camerer (2003) notes that strategic dominance is the most basic principle in decision-making, since it identifies the best strategy unconditional on beliefs about other players' choices.<sup>1</sup> Consequently, the assumption

---

<sup>1</sup>Strategy  $X$  *strictly* dominates strategy  $Y$  if player's payoff from the former is greater than the payoff from the

that all players obey dominance allows for a conservative inference about others' decisions. In some cases, a step-by-step elimination of dominated strategies in the game may yield a unique equilibrium, which renders the game *dominance solvable*. However, as raised by Rosenthal (1981) and Beard and Beil (1994), even as intuitive a solution concept as dominance solvability may not suffice to eliminate suboptimal outcomes in a very simple coordination game. This striking empirical puzzle is the point of departure for Chapters 1 and 2 of the present thesis.

## Public goods games: between the Nash equilibrium and the Pareto efficiency

Public goods game is a decision-making problem where the two important solution criteria, the Nash equilibrium and the Pareto optimality, stand at odds. These problems are often referred to as *social dilemmas*. The simplest and probably the best known illustration of such a problem is a  $2 \times 2$  prisoner's dilemma game. In this stylized problem, the police arrest two criminals suspected of committing a crime. Since there is no sufficient evidence to file charges against them, each of them is put into a separate cell and offered the following bargain: if he testifies against the other person, he will be released and the other person will face a long sentence. Otherwise, he will suffer a short imprisonment himself. Selfishly rational criminals use their dominant strategies and testify against each other, which means a long imprisonment for both. However, if they both cooperate and decide not to testify, they face a shorter sentence. As a result, the unique Nash equilibrium in this game turns out Pareto suboptimal, since it is beneficial for both parties to mutually deviate from the dominant strategy.<sup>2</sup>

Public goods games are useful for capturing important behavioral patterns – *free-riding* and *cooperation* – in strategic contexts related to environmental economics (reducing externalities, sharing a depletable resource) and public economics (contributing to the production of a public utility). Samuelson (1954) notes that free-riding is the main problem with the voluntary provision of public goods, since those who benefit from voluntarily provided public goods have an incentive to avoid paying for them. Subsequent empirical papers in psychology, sociology and economics suggest that although many subjects abstain from free-riding, cooperation is far from being a steadfast norm. For instance, Fischbacher, Gächter, and Fehr (2001) report that one third of players act as typical free-riders, while another half employs a conditionally cooperative strategy – their own willingness to cooperate depends on other players' willingness to cooperate.

An important body of experimental literature finds that institutions may play a crucial role in inducing people to act cooperatively. In particular, numerous studies report that mechanisms based on communication help promote pro-social norms of behavior in economic contexts where individual rationality conflicts with social interest. Chapter 3 of the present thesis is devoted to

---

latter regardless of other players' strategy profile. Strategy  $X$  *weakly* dominates strategy  $Y$  if player's payoff from the former is greater than the payoff from the latter for some strategy profiles utilized by other players, and not lower for all the remaining strategy profiles.

<sup>2</sup>The sequential version of this game is known as *trust game*.

an institution that has been only recently brought to economists' attention as a means of inducing efficiency in social dilemmas: *ex post* communication.

## Information transmission in strategic interactions: theory and experiments

Economists have been holding a long-lasting interest in the question of whether the exchange of information between economic agents affects (and possibly improves) the results of human interactions. Hayek (1945) notes that "[t]he economic problem in the society [...] is a problem of the utilization of knowledge not given to anyone in its totality" (pp. 520-521). In the same spirit, Arrow (1974) identifies internal communication as a key activity of any organization.

### Ex ante communication

The exchange of information is particularly important in those interactive situations that involve strategic uncertainty, *i.e.* exhibit the lack of common knowledge about the behavior of all actors. A simple institution that is often prescribed to improve decision-making in this environment is *cheap-talk communication* which allows agents to signal their intentions to other parties of interaction (Crawford, 1998). Although, *ex vi termini*, this kind of communication is pre-play, costless, non-binding, and non-verifiable for senders and receivers, under certain conditions such messages may be considered credible and affect the decisions of agents whose interests are somewhat aligned (Farrell and Rabin, 1996). For instance, in *coordination games* characterized by the existence of multiple pure-strategy Nash equilibria (which can be furthermore Pareto-rankable), pre-play communication may help players achieve one equilibrium among many others (first-best in the case of Pareto-rankability), and appears as a natural remedy against inefficiencies in coordination problems (Farrell, 1987, 1988). The fundamental game theoretical properties that allow for establishing a link between words and actions are laid out by Farrell and Rabin (1996), Aumann (1990), and Ellingsen and Östling (2010).

Using a standard game theoretical setting with rational decision-makers, Farrell and Rabin introduce the notion of *self-commitment*: if believed by the receiver, a credible message creates incentives for the sender to fulfill it. In particular, in coordination games (like the one studied in Chapters 1 and 2 of the present thesis) cheap-talk communication is self-committing (and thus credible) and makes the Pareto-dominant Nash equilibrium a *self-enforcing agreement* (Farrell, 1988). Aumann points out that whether such agreement would be self-enforcing or not depends on the payoff structure of a coordination game. In particular, communication will not be credible if the sender has a strict preference over the receiver's strategy choice. Aumann argues that in this context messages only convey information on what the sender wants the receiver to do, rather than what the sender intends to do himself. He proposes a more rigorous condition of credibility, dubbed

*self-signaling*: a message is credible if the sender wants it to be believed by the receiver only if it was actually true. Otherwise, a message is uninformative in terms of sender's future behavior and thus should be ignored by the receiver – which would be the case for the coordination game investigated in Chapters 1 and 2. In a recent theoretical contribution, Ellingsen and Östling (2010) depart from the standard rationality assumption and study cheap-talk communication in a level-k bounded rationality framework. Under the central assumption that agents display a weak preference for truthfulness which is of common knowledge, they demonstrate that cheap-talk communication might provide *reassurance* in strategic situations where agents' interests are sufficiently aligned. For instance, in coordination games with Pareto-rankable equilibria these messages convey credible information about sender's rationality – that is, whether he disregards dominated strategies or not.

Experimental evidence suggests that cheap-talk communication is a very powerful institution when it comes to improving the efficiency of human interactions, even in contexts where standard game theory would strongly suggest otherwise.<sup>3</sup> A general consensus in the experimental literature investigating coordination problems is that this kind of cheap-talk communication improves the efficiency of outcomes. Cooper, DeJong, Forsythe, and Ross (1989), Cooper, DeJong, Forsythe, and Ross (1992) and Charness (2000) study various 2-person coordination games (such as the stag hunt game and the battle of sexes game) and find that one-way cheap-talk communication (in which one player acts as the sender and the other as the receiver of a message) substantially enhances efficient coordination. In particular, Charness provides formal evidence that such communication works effectively even when messages are not self-signaling, notwithstanding Aumann's critique. The first two studies also suggest that the mechanics of a two-way communication (in which each agent acts as both the sender and the receiver of messages) is rather subtle: it may outperform one-way communication in certain coordination problems, but lags behind in others. In a subsequent experimental investigation, Clark, Kay, and Sefton (2001) report that the Aumann self-signaling condition plays a central role for the effectiveness of two-way cheap-talk communication. Finally, Blume and Ortmann (2007) find that cheap-talk communication facilitates efficient coordination in large coordination problems that involve more than 2 players (like the minimum-effort game and the median-effort game).

In a broader perspective, the positive behavioral effect of cheap-talk communication may also extend to strategic problems that involve *conflicting interests*. Numerous economic experiments study games in which the Nash equilibrium based on selfish rationality is Pareto-dominated by non-equilibrium solutions requiring cooperative behavior. Data collected by Ellingsen and Johannesson (2004), Charness and Dufwenberg (2006) and Vanberg (2008) for trust games show that promises exchanged prior to decision-making foster trust and trustworthiness, which results in a more

---

<sup>3</sup>It is important to notice at this point that a common methodology in laboratory-based economic experiments (including ours) is to carry out communication using anonymous, written (free-form or fixed-form) messages. For an exhaustive survey on experimental approaches to communication in social science, see Sally (1995).

frequent Pareto-efficient cooperation. Similar evidence is provided by Isaac and Walker (1988) and Bochet, Page, and Putterman (2006) in the context of public goods games. A theoretical approach to these data requires a wide-ranging departure from the standard game theoretical toolbox depicted above. In this vein, Charness and Dufwenberg (2006) turn to *psychological game theory* (which takes into account belief-dependent emotions like reciprocity, guilt aversion, regret and shame) and relate this behavior to *guilt aversion* due to which agents experience disutility from letting down others' payoff expectations. Notwithstanding this explanation, Ellingsen and Johannesson (2004) and Vanberg (2008) argue that the effect of communication is commitment-based: agents display an *aversion to lying* which generates costs of acting inconsistently. This echoes the Brocas, Carrillo, and Dewatripont (2003) argument that a promise may serve as an efficient commitment device if the cost of breaking it is sufficiently high.

## Observation

Rich and encouraging empirical results from the vast literature on cheap-talk communication have given rise to a new strand of experimental research that evaluates the effectiveness of communication relative to another simple, non-monetary institution involving transmission of information: the observation of other players' past behavior. Bracht and Feltovich (2009) and Duffy and Feltovich (2002) report that both institutions help increase the efficiency of interactions, although their effect varies across games. More precisely, observation has a much stronger effect than pre-play communication in situations involving divergent interests (such as a prisoner's dilemma game or a trust game), while the opposite occurs in coordination games where agents' interests are concurrent. In a follow-up study, Duffy and Feltovich (2006) suggest that combining these two institutions may further improve outcomes, but only when multiple signals happen to be aligned (that is, they all point to the same action). Inconsistent signals, in turn, are found to strongly deteriorate the efficiency of outcomes. On the other hand, evidence collected by Wilson and Sell (1997) is far more discouraging. They report that all three mechanisms – communication, observation, and both of them combined – happen to decrease agents' willingness to contribute to the production of a public good as compared to the no-information condition. Chapter 1 offers further experimental evidence on the behavioral effects of communication and observation in human interactions.

## Communication and commitment

While economic literature acknowledges the importance of commitment underlying pre-play communication, this notion still lacks a systematic empirical analysis. To the best of our knowledge, a study by Feltovich and Swierzbinski (2011) is so far the sole experimental investigation of the notion of commitment in communication. In a Nash bargaining game, two kinds of pre-play communication are compared: a classical cheap talk in which messages remain non-binding, and a

protocol where agents may send proposals for the split of the "cake" which, if accepted by the other party, become *binding* agreements. They report that both institutions improve outcomes (higher efficiency, less opting out and less under-demanding), with binding communication furthermore outpacing cheap-talk communication.

Unlike economists, social psychologists have studied in depth the relationship between communication and commitment, especially in the "intermediate states" between the two extreme cases considered by Feltovich and Swierzbinski, *i.e.* a non-binding communication and a fully-binding communication. Kiesler and Sakumura (1966) define commitment as a "binding of the individual to behavioral acts," and Kiesler (1971) further notes that "commitment is a continuous variable rather than a dichotomous one," so that the magnitude of the behavioral effect of committing oneself to a particular task depends on the degree of commitment.

This literature suggests that the force of commitment increases when it is made freely and expressed explicitly – for instance written down and signed, or publicly announced. Evidence from field experiments shows that people who have agreed to sign an undertaking to recycle more paper or save water and electricity become much more devoted to these tasks (Pallack, Cook, and Sullivan, 1980; Wang and Katsev, 1990; Katsev and Wang, 1994; Joule, Girandola, and Bernard, 2007; Guéguen, Joule, Halimi-Falkowicz, Pascual, Fischer-Lokou, and Dufourcq-Brana, 2011).

The strength of commitment may also depend on the form the latter takes. For example, Joule and Beauvois (1998) consider *oath* as a strong commitment device. According to Jacquemet, Joule, Luchini, and Shogren (2013), an oath that is taken freely, expressed publicly and signed, should be considered a stronger commitment device than a verbal promise or a written undertaking. In a recent economic study, they show that making people take an oath to "tell the truth and always provide honest answers" increases their intrinsic motivation to honestly reveal their preferences and helps overcome the hypothetical bias – the so-called gap between hypothetical monetary values and real economic commitments – in an incentive-compatible second-price auction. From this perspective, engaging agents into a truth-telling behavior by means of a voluntary oath appears as a natural way to enhance the commitment underlying cheap-talk communication. Encouraging experimental evidence on communication under oath is provided in Chapter 2.

## **Ex post communication**

Departing even further from the classical game theoretical setup, recent experimental studies investigate communication that seems even cheaper than the one discussed before, since it only occurs *after* decision-making. Interestingly, this communication protocol is found to preserve the important and desirable behavioral properties previously observed in pre-play communication. López-Pérez and Vorsatz (2010) report that the availability of fixed-form, post-play messages makes subjects more cooperative in a prisoner's dilemma game, Ellingsen and Johannesson (2008) and Xiao and Houser (2009) identify the same effect on altruism in subjects playing a dictator



Table 1: Rosenthal’s coordination game

Player A	Player B	
	$l$	$r$
$L$	(9.75;3)	(9.75;3)
$R$	(3;4.75)	(10;5)

game. Czap, Czap, Khachatryan, Burbach, and Lynne (2011) implement a two-stage game in which a common-pool resource is used by a group of subjects, out of which some have private incentives to produce publicly undesirable externalities. They find that the reception of a negative emotional feedback after the first stage reduces externalities in the second stage, while providing positive feedback is detrimental with this respect. Similar evidence is provided by Dugar (2010) for a repeated minimum-effort coordination game: while the availability of negative feedback makes groups of subjects converge towards the Pareto-superior equilibrium, positive feedback pushes the outcomes all the way down to the Pareto-inferior equilibrium. A similar asymmetry between disapproval- and approval-framed feedback is found in a public goods game by Dugar (2013). These authors explain the behavioral effect of post-play communication in the spirit of psychological game theory, pointing to players’s *aversion to being disapproved* by other people. On the other hand, based on experiments involving repeated public goods games Masclet, Noussair, Tucker, and Villeval (2003) and Peeters and Vorsatz (forthcoming) suggest that such communication may be also used to transmit pre-play signals – that is, to express own intentions in the upcoming interactions. Chapter 3 provides an innovative experimental procedure aimed at separating these two potential mechanisms.

## Experimental design

### Rosenthal’s coordination game

Chapters 1 and 2 of the present thesis investigate a coordination game based on Rosenthal (1981) and Beard and Beil (1994). As an illustration, consider the game matrix that is most frequently utilized throughout these chapters and presented in Table 1.

This game exhibits two pure strategy Nash equilibria. A single step of elimination of (weakly) dominated strategies suffices to identify equilibrium  $(R, r)$  which exhibits strong theoretical properties. First, it satisfies the Harsanyi and Selten (1988) criteria of payoff dominance and risk dominance. Second, it is *tremble-hand perfect*, or robust to the risk of accidental off-equilibrium

decisions (Selten, 1975).<sup>4</sup> The other Nash equilibrium,  $(L, l)$ , is Pareto inferior and arises because  $L$  is the action that player A ought to take against a player B choosing (for some reason) a dominated strategy  $l$ . Altogether, in the light of these important game theoretical criteria, the first equilibrium appears to be a natural focal point.

However, accumulated experimental findings – coming from both sequential-move and simultaneous-move versions of the game – robustly show that subjects very often fail to attain the efficient solution. Suboptimal outcomes arise due to two different driving forces. First, players are very frequently reluctant to rely on others’ ability to use dominant strategy, especially when their potential surplus from doing so is not substantial. Second, surprisingly enough, the utilization of dominated strategy indeed happens to be quite widespread. To visualize this puzzle, let us come back to the payoff structure presented in Table 1, where player As are found to abstain from relying on player Bs in more than 50% of interactions, while player Bs maximize their own payoff by selecting a dominant strategy in only 80% of cases. Importantly, these patterns do not vary due to the changing form of interaction (simultaneous-move vs. sequential-move, one-shot vs. repeated). What appears to be even more intriguing, player Bs are also neutral to further variations in the structure of personal payoffs (such as an increased personal cost of using the dominated strategy).

These results not only point to the insufficiency of standard game theory to predict the empirical outcomes of a rather simple collective decision-making task. They also raise an important methodological question of the internal validity of laboratory experiments. In a general perspective, causal inference from laboratory experiments relies on the crucial property that subjects’ behavior is driven by financial incentives, so that *ceteris paribus*, a subject always prefers more money to less. This requirement is key to ensuring that the environment chosen by the experimenter is actually the one in which decisions are made. From both perspectives – theoretical and methodological – a widespread violation of the principle of own-payoff maximization ought to urge experimental economists to identify other key parameters behind economic decisions in humans. First, this paves the way for the future refinements of theoretical models aimed at enhancing their capacity to explain economic phenomena as well as the accuracy of their predictions.<sup>5</sup> Second, it improves the understanding and the reliability of the data delivered by increasingly popular laboratory experiments.

Somewhat in the same vein, Rosenthal himself conjectures that suboptimal outcomes may occur in this game depending on: *(i)* the stakes of the game, and *(ii)* what is known about the interaction partner’s behavior. The first chapter provides a thorough experimental investigation of these conjectures. In the first part, we demonstrate that a well-known model of other-regarding preferences – the inequality aversion model (Fehr and Schmidt, 1999) – offers a powerful theoretical key to explaining unusual patterns in existing data: in fact, the presence of this preference

---

<sup>4</sup>Furthermore, in a sequential version of this game, equilibrium  $(R, r)$  would also be *subgame perfect*.

<sup>5</sup>The idea that not only game theory may inform economic experiments, but also *vice versa*, has been a milestone for a recent development of a new branch of game theory – *behavioral game theory* (see Camerer (2003)).

makes suboptimal outcomes rationalizable for all payoff structures implemented heretofore in lab experiments. Along these lines, we provide an extensive test of the inequality aversion hypothesis in Rosenthal's coordination game. In the second part, we explore the second dimension of his conjecture. We apply a methodology which varies the quantity and the quality of information available to players by allowing subjects to repeat the game, to communicate and to observe others' historical choices prior to taking a decision. The second chapter offers an application of the social psychological theory of commitment to this collective decision-making problem: we test whether an enhanced personal commitment to tell the truth improves the capacity of communication to reduce inefficient behavior.

## The Voluntary Contribution Mechanism

The Voluntary Contribution Mechanism (VCM) is a useful variation of a public goods game which preserves the most important features of the genuine game's equilibrium structure (that is, it exhibits a unique Nash equilibrium which is Pareto-inferior with respect to the non-equilibrium outcomes) and furthermore allows for a flexibility in choosing the size of strategy space and the number of agents (which makes it an important refinement as compared to the most classical representation – the prisoner's dilemma game).

Chapter 3 uses this popular lab implementation of a public goods game to study an institution that has been only recently brought to economists' attention as a means of inducing efficiency in social dilemmas: *ex post* communication. As discussed earlier, existing empirical results are encouraging, but the root underlying the effect of this institution, especially in a repeated interaction, is still rather obscure. This study provides a novel examination of the potential sources of the behavioral impact of *ex post* communication in public goods experiments.

## Overview of results

Chapters 1 and 2 revisit the Beard and Beil (1994) two-player coordination game with two Nash equilibria: one is Pareto-efficient, the other is Pareto-inefficient and involves a weakly dominated strategy. Existing experiments using this game robustly show that suboptimal outcomes arise as a result of two puzzling behaviors: *(i)* subjects doubt that the other players will seek to maximize their own payoff and *(ii)* these doubts are, in some instances, justified.

In Chapter 1, we report on new experiments investigating whether the inequality in payoffs between players, maintained in most lab implementations of this game, may explain such behavior. Our data clearly show that the failure to maximize personal payoffs, as well as the fear that others might act this way, do not stem from inequality aversion. This result is robust to: varying saliency of decisions, repetition-based learning and cultural differences between France and Poland.

Then, we assess whether information about the interaction partner helps eliminate inefficiency

in this game. Our treatments involve three information-enhancing mechanisms: repetition and two kinds of individual signals, messages from partner or observation of his past choices. Repetition-based learning increases the frequencies of the most efficient outcome and the most costly strategic mismatch. Moreover, it is superseded by individual signals. Like previous empirical studies, we report that signals provide a screening of partners' intentions that reduces the frequency of strategic mismatches. Unlike these studies, we find that the transmission of information between partners, either via messages or observation, does not suffice to significantly increase the overall efficiency of outcomes. This happens mostly because additional information does not restrain the use of the dominated action. Therefore, this chapter identifies important limitations of cheap-talk communication – a mechanism generally considered by economists as a useful means to improve the efficiency of economic interactions. It suggests that in the absence of a pronounced link between one's words and actions, institutions involving communication may well happen to be insufficient for this purpose.

This issue is addressed in Chapter 2 where we explore whether the social psychology theory of commitment via a truth-telling oath can improve the performance of pre-play communication regarding the coordination of strategies and the efficiency of outcomes. As an addition to the classical cheap-talk communication protocol utilized in Chapter 1, we ask all players to sign voluntarily a truth-telling oath before entering the lab. Three principle results emerge with commitment-via-the-oath: (1) efficient coordination increases by nearly 50 percent; (2) senders' messages are significantly more truthful and actions more efficient, and (3) receivers' trust of messages increases.

Finally, Chapter 3 turns to *ex post* communication. Using a public goods game based on the voluntary contribution mechanism, it provides a novel experimental testbed for two potential channels by which this mechanism may affect behavior in a repeated interaction: one is related to strategic signaling, the other involves emotions induced by others' judgments. The main findings are as follows. First, the presence of *ex post* communication significantly fosters pro-social behavior. Second, experimental evidence strongly supports the emotion-based hypothesis.

Altogether, this thesis provides a new perspective on the role of communication in economic interactions. As discussed before, economic literature commonly considers communication as a simple way of transmitting signals between rational economic agents who utilize them for attaining their own strategic goals.<sup>6</sup> Notwithstanding this widespread view, evidence reported in this thesis suggests that the nature of communication is not limited to exchanging strategic information. More precisely, we highlight the importance of two psychological factors underlying the effect of communication on human subjects' behavior: personal commitment and aversion to others' expressed disapproval. In this sense, results provided in this thesis not only contribute to a better understanding of the role of information-based institutions in human interactions, but also propose means by which the functioning of these institutions may be improved.

---

<sup>6</sup>See, for instance, a classical study by Crawford and Sobel (1982).

## Résumé

Cette thèse comporte 3 chapîtres principaux. Les chapîtres 1 et 2 présentent des résultats expérimentaux issus d'un jeu de coordination proposé par Rosenthal (1981) et Beard and Beil (1994). Ce jeu comporte deux équilibres de Nash: le premier est efficace, le deuxième repose sur l'usage de stratégies faiblement dominées. Dans les expériences en laboratoire fondées sur ce jeu, les joueurs échouent très souvent à prendre les décisions qui maximisent simultanément les gains de toutes les parties. Ces échecs de coordination efficace proviennent de deux comportements: (i) les sujets doutent que les autres joueurs vont chercher à maximiser leur propre gain, et (ii) ceux doutes sont, dans certains cas, justifiés.

Dans le chapitre 1, nous présentons une nouvelle expérience qui permet de vérifier si ce comportement est dû à par l'inégalité des paiements entre les joueurs (qui subsiste dans la plupart des implémentations de laboratoire menées jusqu'à présent). Nos données montrent clairement que l'échec à maximiser les gains personnels, ainsi que la crainte que les autres pourraient se comporter de cette façon, ne proviennent pas de l'aversion pour l'inégalité. Ce résultat est robuste quant aux variations dans la saillance des décisions, à l'apprentissage par répétition, ainsi qu'aux différences culturelles entre la France et la Pologne.

Nous étudions ensuite l'impact de l'information sur le comportement stratégique dans ce jeu. Les traitements expérimentaux introduisent trois mécanismes améliorant le niveau d'information dans le jeu : une simple répétition, des messages de type "cheap-talk" et l'observation des actions passées du partenaire. L'apprentissage par répétition augmente les fréquences de l'issue la plus efficace, ainsi que le risque de défaut d'appariement stratégique le plus coûteux. De plus, ce type d'apprentissage est remplacé par des signaux individuels. Comme les études précédentes, nous montrons que les signaux aident à prévoir les intentions des partenaires, ce qui réduit la fréquence des échecs de coordination. Néanmoins, contrairement à ces études, nous trouvons que la transmission d'information entre les partenaires, que ce soit en utilisant les messages ou l'observation, ne suffit pas à augmenter significativement l'efficacité globale des résultats. Cela arrive surtout car la transmission d'information ne restreint pas l'utilisation des stratégies dominées.

Dans le chapitre 2, nous proposons une expérience qui applique la théorie de l'engagement, établie en psychologie sociale, dans le contexte économique du jeu de coordination. Dans cet environnement, le jeu de coordination, qui se déroule avec communication, est précédé par l'étape du serment où les sujets ont l'opportunité de s'engager solennellement à dire la vérité. Trois résultats principaux émergent. Tout d'abord, en présence du serment, la coordination sur l'équilibre le plus efficace augmente de près de 50% pour atteindre un niveau de 75%. Ensuite, grâce à la procédure du serment, les joueurs deviennent plus honnêtes: ils envoient des messages qui correspondent plus souvent à ce qu'ils font effectivement dans le jeu. De plus, les actions qu'ils choisissent sont aussi plus efficaces. Enfin, les joueurs qui reçoivent les messages, deviennent plus confiants et ils choisissent plus souvent une action conforme aux intentions qui leur sont envoyées.

Le troisième chapitre, c'est le groups des études expérimentales qui montrent que la communication *a posteriori* promeut la générosité dans les dilemmes sociaux où les incitations individuelles sont en contradiction avec l'intérêt commun, comme la contribution aux biens publics. Néanmoins, la nature de cette institution, notamment dans une interaction répétée, reste largement inexpliquée. Mon étude propose un test empirique de deux mécanismes par lesquels la communication *a posteriori* peut influencer le comportement dans les interactions répétées: l'un est lié à la signalisation stratégique, l'autre implique des émotions induit par l'opinion des autres. La présence de la communication *a posteriori* (menée par l'attribution de points de désapprobation gratuits) renforce le comportement pro-social et réduit le free-riding. Nous trouvons des preuves systématiques en faveur de l'hypothèse basée sur l'émotion. Cependant, nous ne trouvons aucun support statistique pour l'hypothèse basée sur la signalisation. Une interprétation possible de ce phénomène est que les messages consistent une punition non-monétaire face aux passagers clandestins.

# Chapter 1

## Inequality aversion and information transmission in dominance-solvable coordination games

*Based on joint work with Nicolas Jacquemet.<sup>1</sup>*

### 1.1 Introduction

In coordination games, people may fail to synchronize their actions because of strategic uncertainty – the risk associated with not knowing how other players will play the game. In such contexts giving rise to multiple equilibria, theoretical refinements are characterized by assumptions on players' beliefs about other players' behavior. For instance, the idea that players rely on the own-payoff-maximizing behavior of others is enough in theory to rule out equilibria supported by non-credible strategies that may undermine the value of interaction.

Rosenthal (1981) challenges this view in the context of a sequential, two-player, one-shot game in which the first mover either decides alone on the final issue of the game or relies on the second mover. In the latter case, second mover only has to decide whether to maximize both players' payoffs or not. But since the first mover stands to lose a lot should the second mover fail to maximize payoffs, a very likely outcome is the Pareto-dominated one in which the first mover decides alone. For this reason, this game is one of the simplest textbook example of the possible failure of subgame perfectness. Rosenthal conjectures that an imperfect equilibrium may occur depending on: *(i)* the stakes of the game, and *(ii)* what is known about the interaction partner's behavior.

---

<sup>1</sup>Material presented in Section 1.2 is based on Jacquemet and Zylbersztejn (2011), while Section 1.3 is built upon Jacquemet and Zylbersztejn (2012).

The accumulated experimental evidence coming from both sequential-move and simultaneous-move versions of the game – strongly supports the first part of the Rosenthal conjecture. In these lab implementations (overviewed in 1.1), suboptimal outcomes arise due to two different driving forces. First, as expected by Rosenthal, subjects who play as first movers are very frequently reluctant to rely on second movers’ ability to maximize own payoff, especially when their potential surplus from doing so is not substantial. Second, surprisingly enough, an important share of second movers do indeed turn out to be unreliable and fail to maximize both players’ payoffs.

The second part of this chapter (Section 1.2) revisits Rosenthal’s first conjecture and offers a new perspective to look at existing empirical data. In line with recent developments in behavioral economics and game theory, we depart from the standard decision-making framework that is based on a selfishly rational *homo oeconomicus* by incorporating other-regarding preferences – namely, aversion to inequality (Fehr and Schmidt, 1999). *Ex ante*, this model offers a powerful theoretical key to explaining unusual patterns in existing data: in fact, the presence of aversion to inequality makes suboptimal outcomes rationalizable for all payoff structures implemented heretofore in lab experiments. Finally, Section 1.4 concludes.

In the third part of this chapter (Section 1.2), we provide a seminal investigation of the second dimension of the Rosenthal conjecture – the importance of the availability of information about the interaction partner. More precisely, we address the question of whether and how information-enhancing institutions can help overcome coordination failures for a *given* payoff structure.

## Two puzzling behaviors in experiments using Rosenthal’s game

Table 2.2 provides an overview of previous experimental implementations of the game studied in this paper. In all cases, there are only three different outcomes in the game, detailed in the left-hand side of the table: either player A chooses to decide alone by picking  $L$ , or he relies on player B’s decision by choosing  $R$ . In this case, both players’ payoffs are higher if  $r$  is chosen rather than  $l$ . The right-hand side of the table summarizes each outcome as a percentage of all observed decisions, as well as the frequency of action  $r$  conditional on reliance from player A.

The main focus of the original study by Beard and Beil (1994) is to test Rosenthal’s conjecture that subjects may be unwilling to rely on the other players’ ability to maximize payoffs – hence challenging subgame perfectness. Their experimental evidence supports the Rosenthal conjecture: although the share of payoff-maximizing subjects is high in the sub-population of player Bs who are relied upon (from 83% in their treatment 1 to 100% in treatments 3 and 4, for instance), most player As decide not to rely on their partners. The comparison across treatments shows that behavior highly depends on the size of stakes: the lower the value of the secure option, the higher the reliance rate from player As; the higher the cost of being unreliable, the higher the reliability rate from player Bs. Beard, Beil, and Mataga (2001) replicate some of these treatments using Japanese subjects. While the results of their treatment 1 are in line with previous US evidence, behavior



Table 1.1: Summary of experimental evidence on Rosenthal’s game

Experiment	Form	Payoff			Outcomes (%)					Nb. obs.
		( $L$ )	( $R, r$ )	( $R, l$ )	$L$	$R, r$	$R, l$	$r R$	$r$	
Beard, Beil–Tr.1	Seq	(9.75; 3)	(10; 5)	(3; 4.75)	66	29	6	83	—	35
Beard, Beil–Tr.3	Seq	(7.00; 3)	(10; 5)	(3; 4.75)	20	80	0	100	—	25
Beard, Beil–Tr.4	Seq	(9.75; 3)	(10; 5)	(3; 3)	47	53	0	100	—	32
Beard et al.–Tr.1	Seq	(1450; 450)	(1500; 750)	(450; 700)	79	18	3	83	—	34
Beard et al.–Tr.2	Seq	(1050; 450)	(1500; 750)	(450; 700)	50	18	32	64	—	28
Goeree, Holt–Tr.1	Ext	(80; 50)	(90; 70)	(20; 10)	16	84	0	100	—	25
Goeree, Holt–Tr.2	Ext	(80; 50)	(90; 70)	(20; 68)	52	36	12	75	—	25
Goeree, Holt–Tr.3	Ext	(400; 250)	(450; 350)	(100; 348)	80	16	4	80	—	25
Cooper, Van Huyck–Tr.9	Str	(4; 1)	(6; 5)	(2; 4)	27	—	—	—	86	187
Cooper, Van Huyck–Tr.9	Ext	(4; 1)	(6; 5)	(2; 4)	21	—	—	—	84	187

**Note.** Several representations of the game have been applied so far, as stated in column 1: simultaneous-move strategic-form game (Str), simultaneous-move extensive-form game (Ext), sequential-move game (Seq). The monetary payoffs displayed in columns 2-4 are in USD in Beard & Beil (1994) and Cooper and Van Huyck (2003), in US cents in Goeree & Holt (2001), in yens in Beard et al. (2001) and in euros in our treatments.

in treatment 2 is in a sense more striking. Because of the increase in the monetary incentives to reach  $(R, r)$  as compared with the secure outcome  $(L)$ , a much higher share of player As rely on their partner – from 21% in treatment 1 to 50% in treatment 2. However, although the payoffs for player B remain unchanged between the two treatments, the share of payoff-maximizing decisions (conditional on having been relied upon in the first place) falls from 83% to 64%. Importantly, both of these studies elicit decisions from player Bs conditional on the prior reliance of player As. Hence, this evidence does not suffice to infer the actual behavior in the entire population of player Bs, since the decisions of player Bs paired with unreliable player As cannot be observed.

In contrast, Goeree and Holt (2001) and Cooper and Van Huyck (2003) use a simultaneous-move game of either normal or extensive form, and confirm the robustness of previous evidence. First, even when player Bs are actually reliable (such as in the Goeree and Holt, 2001 treatment 1), a large share of player As are reluctant to rely on them. Second, in a large variety of treatments (treatments 2 and 3 in Goeree and Holt, 2001 and both treatments of Cooper and Van Huyck, 2003) player Bs actually are often unreliable – decision  $r$  occurs from 15% to 25% of the time.

## 1.2 Inequality aversion and efficiency in coordination games

In this section, we explore whether the source of inefficient behavior described in Section 1.1 may be related to other-regarding preferences – namely, the aversion to inequality. The extent to which this preference explains the divergence between human decisions and standard game-theoretical predictions is the subject of lively debate in experimental economics. For instance, lab experiments

Table 1.2: Generic game

player B		
player A	$l$	$r$
$L$	$(a;b)$	$(a;b)$
$R$	$(a-c;b-e)$	$(a+d;b+f)$

by Charness and Rabin (2002) and Engelmann and Strobel (2004) provide evidence against the inequality aversion hypothesis, while subsequent experiments by Fehr, Naef, and Schmidt (2006), Bolton and Ockenfels (2006), Blanco, Engelmann, and Normann (2011), as well as a neuroeconomic study by Tricomi, Rangel, Camerer, and O'Doherty (2010), report evidence in its favor. To the best of my knowledge, the present study is the first attempt to explore the relationship between equality-oriented preferences and the occurrence of inefficiencies in coordination games. This investigation is based on two research questions: do unequal payoffs generated by the dominant outcome *(i)* make second movers unwilling to maximize both players' payoffs, and *(ii)* make first movers reluctant to rely on actually reliable second movers? We implement six variations of the payoff structure of the game, all sharing the main strategic dimensions of the original formulation. The Baselines are compared to Egalitarian Treatments, which restore the equality between players in the payoffs generated by the Pareto-dominant outcome. We also provide evidence of the cross-cultural robustness of these results by collecting data in France and in Poland.

### 1.2.1 Empirical strategy

From existing experimental evidence (summarized in Section 1.1), two puzzling behaviors arise: a large share of player As appear reluctant to rely on actually reliable player Bs; but at the same time the weakly-dominated strategy is often chosen by player Bs. The purpose of this paper is to test whether such behavior is related to the payoff structure of the game.<sup>2</sup> To illustrate the point, the general structure of the game is presented in Table 1.2 – of which all the lab implementations presented in Table 2.2 above are particular cases. The crucial properties of the game hold if the parameters  $a, \dots, f$  take non-negative values and all payoffs are positive. Player B (weakly) maximizes both players' payoffs by selecting  $r$ , while player A prefers  $L$  to the outcome  $(R, l)$  that

<sup>2</sup>This hypothesis has been already raised in the literature – see for instance (Goeree and Holt, 2001, p.1416) – but to the best of our knowledge it has never been examined empirically. One exception is treatment 6 in Beard and Beil (1994), discussed in Section 1.2.1. Surprisingly, this treatment is not commented on in the original paper, neither it is discussed as a means to assess the sensitivity of behavior to more equalized payoffs. In any case, as stressed above, the original design of Beard and Beil (1994) is inappropriate for studying player Bs' behavior, since their decisions are elicited only conditional on player A's choice.

is attained when relying on an unreliable player B. As a result, the game has two Nash equilibria,  $(L, l)$  and  $(R, r)$ . The first is imperfect (since it entails player B's weakly dominant strategy  $l$ ), while the second is (trembling-hand) perfect and Pareto-superior.

But in most experimental implementations of this game, the payoffs from the perfect Nash equilibrium are much higher for player A than for player B, *i.e.*,  $a \geq b$ ,  $a + d \geq b + f$ ,  $b - e \geq a - c$ . Although this does not make B's unreliable decision  $l$  a rational answer to A's reliance, non-standard preferences involving aversion to inequality might explain why player Bs forgo efficiency at a personal monetary cost Fehr and Schmidt (1999). Moreover, if player As believe that this sort of preference exists among player Bs, they may prefer to choose the secure option instead of relying on their partner. In order to test this hypothesis, our empirical approach consists in complementing the original payoff structure with variations that maintain the strategic properties of the game, while equalizing the payoffs both players earn in the case of payoff-maximizing coordination, *i.e.*, such that  $(a + d) = (b + f)$ .

The treatment effects we seek to identify are best illustrated in the framework of the inequality aversion model by Fehr and Schmidt (1999). Both subjects  $i, j \in \{A, B\}$  are assumed to choose their actions in the game presented in Table 1.2 according to the extended utility function defined on outcome  $O$  generating payoffs  $(O_i; O_j)$ :

$$U_i(O|\alpha_i, \beta_i) = O_i - \alpha_i * (O_j - O_i) * \mathbf{1}_{O_i < O_j} - \beta_i * (O_i - O_j) * \mathbf{1}_{O_i > O_j} \quad (1.1)$$

Parameters  $0 \leq \beta_i \leq \alpha_i$  measure the sensitivity of player  $i$  to inequality ( $\mathbf{1}_{O_i < O_j} = 1 - \mathbf{1}_{O_i > O_j} = 1$  if  $j$  earns more than  $i$ , 0 otherwise). The payoff structures of our Baseline Treatments (BT1 and BT2 in Table 1.3) are such that  $\exists(\alpha_B, \beta_B) : U_B((R, l)^{BT}|\alpha_B, \beta_B) > U_B((R, r)^{BT}|\alpha_B, \beta_B)$ , so that a player B whose utility is defined by (1.1) may prefer outcome  $(R, l)$  over  $(R, r)$ . By the same token, Egalitarian Treatments (ET1-4 in Table 1.3) are built in such a way that  $\forall(\alpha_B, \beta_B) : U_B((R, l)^{ET}|\alpha_B, \beta_B) < U_B((R, r)^{ET}|\alpha_B, \beta_B)$ , so that for the same preference parameters, player B now prefers  $(R, r)$  over  $(R, l)$ . As for player As, define  $\theta_T$  as the perceived likelihood that player Bs' realization of  $(\alpha_B, \beta_B)$  makes him prefer  $(R, l)$  over  $(R, r)$  in Treatment  $T$ . If player A is a risk-neutral expected utility maximizer, then  $EU_A(R^T|\alpha_A, \beta_A, \theta_T) = \theta_T U_A((R, l)^T|\alpha_A, \beta_A) + (1 - \theta_T) U_A((R, r)^T|\alpha_A, \beta_A)$ . As a result, if the fear that player Bs are inequality averse is high enough, player A may prefer the secure choice in the Baseline Treatment, since  $\exists(\alpha_A, \beta_A, \theta_{BT}) : U_A(L^{BT}|\alpha_A, \beta_A) \geq EU_A(R^{BT}|\alpha_A, \beta_A, \theta_{BT})$ . By contrast, the Egalitarian Treatment is designed in such a way that  $\theta_{ET} = 0$ .

## Overview of the experimental design

Table 1.3 provides a full overview of the payoff variations we implement. Our starting point is Baseline Treatment 1 – which corresponds to the Beard and Beil (1994) treatment 1 and the its replication used as baseline in the previous section – with which we associate an egalitarian

Table 1.3: Overview of experimental treatments

Treatments		player A chooses $L$	player A chooses $R$		Nb. of Sessions Paris    Warsaw		Likelihood of decisions	
			player B chooses $l$	player B chooses $r$			$R$	$r$
Baseline 1	(BT1)	(9.75 ; 3.00)	(3.00 ; 4.75)	(10.0 ; 5.0)	3	–	0.490	0.807
Egalitarian 1	(ET1)	(9.75 ; 5.00)	(5.00 ; 9.75)	(10.0 ; 10.0)	3	–	0.457	0.727
Egalitarian 3	(ET3)	(9.75 ; 5.50)	(5.50 ; 8.50)	(10.0 ; 10.0)	2	2	0.575	0.828
Egalitarian 4	(ET4)	(8.50 ; 5.50)	(5.50 ; 8.50)	(10.0 ; 10.0)	2	1	0.730	0.823
Egalitarian 2	(ET2)	(8.50 ; 8.50)	(6.50 ; 8.50)	(10.0 ; 10.0)	3	2	0.776	0.936
Baseline 2	(BT2)	(8.50 ; 7.00)	(6.50 ; 7.00)	(10.0 ; 8.5)	3	–	0.743	0.940

**Note.** For each treatment, the first three columns provide the payoffs of each player, in euros. The next two columns give the number of 20 subjects-sessions run in each location. The last two columns summarize the average behavior observed in each treatment over all subjects and repetitions.

version, Egalitarian Treatment 1 – this is also treatment 6 in Beard and Beil (1994). As shown in the last two columns of the table, presenting the observed proportions of reliant decisions ( $R$ ) from player As and reliable decisions ( $r$ ) from player Bs, equalizing payoffs in the Pareto-efficient outcome has virtually no effect on players’ behavior. An important shortcoming of this pair of treatments is that saliency may be violated, as reliance and reliability only induce a 0.25 euro variation in subjects’ payoffs.

In Egalitarian Treatment 3, built on ET1, we improve the saliency of reliable decisions for player B: the opportunity cost of playing  $r$  (instead of the weakly dominated decision  $l$ ) against  $R$  is now 1.5 euros.<sup>3</sup> Similarly, in Egalitarian Treatment 4 based on ET3, we improve the saliency of being reliant for player A – the payoff difference between the two Nash equilibria amounts to 1.5 euros instead of 0.25 euros. In both of these treatments, we maintain equality in payoffs between players in the Pareto-efficient outcome. We derive mixed results from these two additional treatments: the behavior of player Bs is unaffected by both changes, while the share of reliant player As significantly increases (from around 50% to 73%) between ET3 and ET4.

In all treatments described up till now, a strong payoff inequality between players remains if player A chooses  $L$ . As a result, it could be that A’s choice of being reliant is itself driven by the willingness to move away from outcomes with too much inequality in payoffs. In Egalitarian Treatment 2, built on ET4, we restore equality in the imperfect equilibrium by raising player B’s payoff.<sup>4</sup> Surprisingly enough, the main outcome from this treatment is that it strongly disciplines the behavior of player Bs – the weakly dominated action is now almost never chosen. The behavior

<sup>3</sup>In the course of trials leading to the current experimental treatments, we also slightly raised all payoffs from 5.0 in ET1 to 5.5 in ET3, to verify whether decimals have any effect on players’ behavior.

<sup>4</sup>By the same token as above, we slightly raise the payoff earned by player A in the event of unsuccessful attempts to rely on B, from 5.5 to 6.5.

of player As, by contrast, is unaffected by this change. To assess whether such a high rate of secure choices – despite widespread reliability in the population of player Bs – is reinforced by payoff inequality, we complete our design with Baseline Treatment 2, built on ET2, in which the Pareto-dominant outcome gives rise to payoff inequality in favor of player A.<sup>5</sup> Again, we fail to find any effect of this dimension of the game, as both players behave in the same manner as in ET2.

All our main treatments of interest were run in Paris (France). The last concern we want to address is the possibility that our treatment effects are driven by culture-specific motives. We provide robustness evidence on this dimension through additional sessions run in Warsaw (Poland). We chose the three treatments in which observed behavior is in our view the most striking, *i.e.* ET2, ET3 and ET4. The number of sessions in each location is provided in Table 1.3.

## Experimental procedures

Each payoff variation described above defines a treatment, and all treatments are implemented separately using a between-subject design. The design is kept the same for all treatments.

We introduce two important changes to Beard and Beil’s original design. First, we allow for learning by repeating the one-shot game 10 times.<sup>6</sup> Each occurrence is one-shot in the sense that: roles are fixed; pairs are re-matched in each round using a perfect-stranger, round-robin procedure;<sup>7</sup> we avoid the end-game effect by providing no information about the exact number of repetitions; and take-home earnings are derived from one round, randomly drawn out of the ten at the end of each experimental session. Second, we elicit both players’ decisions in each occurrence of the game. In that respect, we break the original sequentiality of the game and ask each player for unconditional choices in each round. Players are only informed about their own payoffs at the end of each round.

A typical session proceeds as follows. Upon arrival, subjects sign an individual consent form. They then enter the lab, where they are randomly assigned to their computers and asked to fill in a small personal questionnaire containing basic questions about their age, gender, education, etc. The written instructions are then read aloud.<sup>8</sup> Players are informed that they will play

---

<sup>5</sup>In designing this treatment, we seek to introduce payoff inequality between players, while holding constant the saliency of being reliable for player B. We thus choose to reduce player B’s payoff in  $(R, l)$  to 7.00, instead of 8.50 in ET2, and accordingly adjust the payoff stemming from decision  $L$ .

<sup>6</sup>Although Rosenthal made his conjecture for a one-shot game, Beard and Beil note in their paper (pp. 261-262) that it seems equally valid for repeated play. The authors furthermore state that learning through experience may affect people’s behavior independently of payoff-related factors.

<sup>7</sup>This procedure, also known as rotation matching, is optimal for our experimental design: for a given number of players and the one-shot nature of each interaction between subjects, it maximizes the number of rounds. See Kamecke (1997) and Duffy and Feltovich (2002) for a related discussion.

<sup>8</sup>An English translation of the original instructions written in French or Polish is provided as supplementary material in Section 1.5.

some (unrevealed) number of rounds of the same game, each round with a different partner, and that their own role will not change during the experiment. Before starting, subjects are asked to fill in a quiz assessing their understanding of the game they are about to play. Once the quiz and all remaining questions have been answered, the experiment begins. Prior to the first round, players are randomly assigned to their roles – either A or B. They are then anonymously and randomly matched to a partner and asked for their choice:  $R$  or  $L$  for player As,  $r$  or  $l$  for player Bs. At the end of each round, each player is informed only about their own payoff. When all pairs have completed a round of the game, the subjects are informed whether or not a new round is starting. In the event of a new round, pairs are re-matched according to a perfect-stranger matching procedure (any pair meets only once in the session). At the end of the experiment, one round is randomly drawn and each player receives the amount corresponding to their gains in that round plus a show-up fee.

All in all, we ran a total of 21 experimental sessions: 16 in the Laboratoire d'Economie Experimentale de Paris (LEEP) at University Paris 1 Panthéon-Sorbonne, between June 2009 and January 2012, and 5 in the Experimental Economics Laboratory at Warsaw University in February 2012.<sup>9</sup> In both laboratories, subjects were recruited via an online registration interface adapted using ORSEE (Greiner, 2004) and the experiment was computerized using software developed under REGATE (Zeiliger, 2000). Each session lasted about 45 minutes, with an average payoff of 14 euros in Paris and 28 PLN (about 7 euros) in Warsaw.<sup>10</sup> No subject participated in more than one experimental session.

Out of 420 participants, 206 were men and 214 were women. A vast majority of the population (332 subjects) were students in various disciplines, and 193 of them were likely to have some knowledge of game theory due to their field of study.<sup>11</sup> 297 subjects had taken part in economic experiments before. The average age of the participants was 24.

### 1.2.2 Statistical procedure for mean comparisons

The experimental design raises the issue of two kinds of correlation in the data. First, since players make a sequence of decisions, each subject's choices might be serially correlated. Second, interaction partners change after each round of the experiment, which might result in an inter-subject correlation. To account for this structure of the data, we perform statistical tests for the

---

<sup>9</sup>Data from the sixth session run in Warsaw, implementing ET4, has been lost as a result of a software crash.

<sup>10</sup>For French subjects, all the payoffs in experimental instructions were expressed in euros. For Polish subjects, we use the same payoff scheme, but expressed in Experimental Current Units (ECU). For the purposes of payment, ECU were converted to Polish Zloty (PLN) at the rate 1 ECU=2 PLN. The participation fee was 5 euros in Paris and 10 PLN (around 2.5 euros according to the current exchange rate in 2012) in Warsaw. Since a vast majority of our subjects are students, and petty student jobs usually pay about 8 euros in Paris and 15 PLN in Warsaw, we strongly believe that the participants' monetary incentives are comparable between countries.

<sup>11</sup>Disciplines such as economics, engineering, management, political science, psychology, applied mathematics for the social sciences, mathematics, computer science, and sociology.

comparisons of means through parametric regressions that assume clustered standard errors at the session level. This specification is asymptotically robust to any misspecification of the OLS residuals (Williams, 2000; Wooldridge, 2003). We moreover apply a delete-one jackknife correction in order to account for a potential small sample bias. Note, observations in the first round are still independent within and between sessions. Thus, we use two-sided Fisher's exact test for mean comparisons in round 1.

### Standard errors estimation

The data are split into clusters (at the session level) and we denote  $i$  each observation in each cluster  $s$ , with  $i=\{1, \dots, N_s\}$  and  $s=\{1, \dots, S\}$  so that the total number of observations is  $N = \sum_{s=1}^S N_s$ . We perform statistical tests for differences between means through linear probability models of the form:

$$y_{is} = \sum_{k=0}^K \beta_k x_{is,k} + \epsilon_{is}$$

in which  $y_{is}$  is a dummy dependent variable,  $X_{is} = \{1, x_{is1}, \dots, x_{isK}\}$  is the set of explanatory variables including the intercept,  $\{\beta_0, \dots, \beta_K\}$  is the set of unknown parameters, and  $\epsilon_{is}$  is the error term. We will consider regressions on dummy variables reflecting changes in the environment (for instance, experimental treatments). Because the endogenous variable is itself binary, we moreover have that:  $E(y|X) = Pr(y = 1|X)$ . In this specification, the parameters thus reflect the mean change in the probability of the outcome induced by the change in the environment. In the case of one explanatory variable,  $y_{is} = \beta_0 + \beta_1 I_{is} + \epsilon_{is}$ , for instance:  $E(y_{is}|I_{is} = 1) - E(y_{is}|I_{is} = 0) = Pr(y_{is} = 1|I_{is} = 1) - Pr(y_{is} = 1|I_{is} = 0) = \beta_1$ , so that the parameter measures the mean variation in the probability of  $y$ . Two-sided  $t$ -tests on each treatment-related parameter thus provide significance levels of the differences in means with respect to baseline condition.

To compute the standard errors, we allow for dependences inside clusters as well as unspecified heteroscedasticity across observations,<sup>12</sup> *i.e.* we assume that any two error terms  $i$  and  $j$  are independent between clusters,  $Cov(\epsilon_{ig}, \epsilon_{jh}) = 0 \forall g \neq h$ , but allow for any type of dependence within a cluster,  $Cov(\epsilon_{ig}, \epsilon_{jg}) = \sigma_{ijg}^2 \forall i, j, g$ . To that end, we correct the estimated covariance matrix at the cluster level using the following procedure, in which the model is written at the cluster level,  $Y_s = X_s \beta + \epsilon_s$ , where  $Y_s$  and  $\epsilon_s$  are  $[N_s \times 1]$  vectors,  $X_s$  is a  $[N_s \times (K+1)]$  matrix,  $\beta$  is a  $[(K+1) \times 1]$  vector:

1. Using the parameters estimated on pooled data,  $\hat{\beta}_{OLS} = (X'X)^{-1}(X'Y)$ , we calculate the vector of error terms in each cluster:

---

<sup>12</sup>Heteroscedasticity is due to the linear probability specification. Even if the data generating process was *i.i.d* (*i.e.*  $V(u_{is}) = \sigma^2$ , and  $E(u_{is}u_{jt}) = 0 \forall i \neq j$  and  $\forall t$ ) the model implies that:  $V(y|X) = Pr(y = 1|X)[1 - Pr(y = 1|X)] = X\beta(1 - X\beta)$ .

$$\hat{\epsilon}_s = Y_s - X_s \hat{\beta}_{OLS}$$

2. We then estimate the cluster robust covariance matrix (CRCME):

$$\hat{V}_{CRCME} = (X'X)^{-1} \left( \sum_{s=1}^S X'_s \hat{\epsilon}_s \hat{\epsilon}'_s X_s \right) (X'X)^{-1} \quad (1.2)$$

### Correction for small sample bias in standard errors

The procedure described above provides a consistent estimator of the covariance matrix which can typically be biased in small samples. What is more, the bias is generally found to be negative, so that significance tests reject the null hypothesis too often. A first way to deal with this issue is to correct for the degrees of freedom by substituting  $\tilde{\epsilon}_s = \sqrt{C_{df}} \hat{\epsilon}_s$ , with  $C_{df} = \frac{S(N-1)}{(S-1)(N-K)}$ , in (1.2) – a procedure known in the literature as HC1. Bell and McCaffrey (2002); Cameron, Gelbach, and Miller (2008) propose a more accurate correction, called HC3, which estimates the residuals as  $\tilde{\epsilon}_s = \sqrt{\frac{S-1}{S}} [I_{N_s} - H_{ss}]^{-1} \hat{\epsilon}_s$ , where  $I_{N_s}$  is a  $[N_s \times N_s]$  identity matrix, and  $H_{ss} = X_s (X'X)^{-1} X'_s$ . For an OLS regression, this corrected variance-covariance matrix amounts to implement a delete-one jackknife procedure:

$$\tilde{V}_{jackknife} = \frac{S-1}{S} \sum_{s=1}^S \left( \tilde{\beta}_{-s} - \hat{\beta} \right) \left( \tilde{\beta}_{-s} - \hat{\beta} \right)' \quad (1.3)$$

where  $\tilde{\beta}_{-s}$  is the vector of coefficients estimated after leaving out the  $s$ th cluster.<sup>13</sup>

### 1.2.3 Results

To ease the description of observed behavior, we use the following nomenclature: a *cooperative decision* means  $R$  ( $r$ ) for player As (Bs). A *cooperative outcome* accordingly corresponds to the Pareto-Nash equilibrium  $(R, r)$ , while a *coordinated outcome* describes a situation where either of the existing Nash equilibria –  $(L, l)$  or  $(R, r)$  – is attained. Last, we label coordination failures following a statistical terminology, where *Type 1 error* corresponds to outcome  $(L, r)$ , while  $(R, l)$  is a *Type 2 error*.

---

<sup>13</sup>All  $p$ -values presented in the section below are associated to statistics computed according to this HC3 procedure. We also ran robustness checks by implementing the HC1 correction, which generally leads to lower estimated standard errors. Our choice is thus conservative as regards our ability to find significant differences in behavior. Based on a correction closely related to the HC3 procedure, Angrist and Lavy (2009) find an inflation of the cluster-robust standard errors by 10% up to 50%.



Table 1.4: Observed behavior in Treatments 1

Game		N	Decisions			Coordination		Failure	
			$R$	$r$	$r R$	$(R, r)$	$(L, l)$	$(L, r)$	$(R, l)$
Beard and Beil (1994, Tr.1)		35	0.343	—	0.833	0.286	—	—	0.057
Baseline Treatment 1	Round 1	30	0.233	0.800	1.000	0.233	0.200	0.567	0.000
	Rounds 2-10	270	0.519	0.807	0.836	0.433	0.107	0.374	0.085
	Overall	300	0.490	0.807	0.844	0.413	0.117	0.393	0.077
Beard and Beil (1994, Tr. 6)		26	0.692	—	1.000	0.692	—	—	0.000
Egalitarian Treatment 1	Round 1	30	0.467	0.633	0.643	0.300	0.200	0.333	0.167
	Rounds 2-10	270	0.456	0.737	0.732	0.333	0.141	0.404	0.122
	Overall	300	0.457	0.727	0.723	0.330	0.147	0.400	0.127

**Note.** For each treatment (in rows), the first column gives the total number of observations – the number of subjects is equal to  $N$  in Beard and Beil’s experiment, and  $N/10$  in ours. The three subsequent columns present the unconditional decisions observed in each treatment and the rate of reliability conditional on reliance from player A (relevant especially for Beard and Beil’s data). The last four columns present the empirical frequencies of the four possible outcomes (two of which –  $(L, l)$  and  $(L, r)$  – are not observable in Beard and Beil’s data).

### Replication of Beard and Beil’s treatments

Our first comparison of interest is BT1 against ET1, two treatments previously studied by Beard and Beil (1994). In Table 1.4, we summarize the main outcomes from both their and our implementations of these two treatments. The data collected by Beard and Beil (1994) in the context of a one-shot sequential-move game seem to favor the inequality aversion hypothesis. The rate of reliance by player As doubles, from 34.3% (12 cases out of 35) in the Baseline Treatment 1 to 69.3% (18/26) in the Egalitarian Treatment 1 ( $p=0.010$ ), and the rate of conditional reliability from player Bs increases from 83.3% (10/12) to 100% (18/18) ( $p=0.152$ ). Therefore, player As do indeed act more cooperatively in an environment where bilateral cooperation gives equal profits to both partners, and player Bs actually seem to prefer an equal division of the gains for cooperation, although the magnitude of this preference is marginal.

Our replication of both treatments in a repeated simultaneous-move game yields results that are more mixed and – most importantly – non-persistent. For inexperienced player As (actions observed in round 1), the likelihood of action  $R$  also doubles, as in Beard and Beil’s study, increasing from 23.3% in BT1 to 46.7% in ET1 ( $p=0.103$ ), but player Bs are at the same time marginally less reliable – the frequency of cooperative choices falls from 80% to 63.3% ( $p=0.252$ ). Importantly, these differences completely disappear among experienced players (actions observed in rounds 2-10), where subjects’ behavior is virtually identical under both conditions: the frequency of decisions  $R$  varies between 51.9% in BT1 and 45.6% in ET1 ( $p=0.574$ ), while the frequencies of decisions  $r$  by player Bs are 80.7% and 73.7% ( $p=0.602$ ).<sup>14</sup>

<sup>14</sup>In line with individual behavior, outcomes do not react much to treatment. Overall, only 53% of outcomes in

Table 1.5: Observed decisions

		N	Round 1			Rounds 2-10			Overall		
			$R$	$r$	$r R$	$R$	$r$	$r R$	$R$	$r$	$r R$
Baseline Treatment 1		300	<b>0.233</b>	<b>0.800</b>	<b>1.000</b>	<b>0.519</b>	<b>0.807</b>	<b>0.836</b>	<b>0.490</b>	<b>0.807</b>	<b>0.844</b>
Egalitarian Treatment 1		300	<b>0.467</b>	<b>0.633</b>	<b>0.643</b>	<b>0.456</b>	<b>0.737</b>	<b>0.732</b>	<b>0.457</b>	<b>0.727</b>	<b>0.723</b>
Egalitarian Treatment 3	France	200	0.600	0.750	0.917	0.611	0.761	0.773	0.610	0.760	0.787
	Poland	200	0.550	0.850	0.909	0.539	0.900	0.900	0.540	0.895	0.898
	Overall	400	<b>0.575</b>	<b>0.800</b>	<b>0.913</b>	<b>0.575</b>	<b>0.831</b>	<b>0.831</b>	<b>0.575</b>	<b>0.828</b>	<b>0.839</b>
Egalitarian Treatment 4	France	200	0.650	0.700	0.615	0.756	0.833	0.838	0.745	0.820	0.819
	Poland	100	0.600	0.900	1.000	0.711	0.822	0.828	0.700	0.830	0.843
	Overall	300	<b>0.633</b>	<b>0.767</b>	<b>0.737</b>	<b>0.741</b>	<b>0.830</b>	<b>0.835</b>	<b>0.730</b>	<b>0.823</b>	<b>0.826</b>
Egalitarian Treatment 2	France	300	0.500	0.867	0.867	0.766	0.952	0.947	0.740	0.943	0.941
	Poland	200	0.600	0.800	0.667	0.856	0.939	0.935	0.830	0.925	0.916
	Overall	500	<b>0.540</b>	<b>0.840</b>	<b>0.778</b>	<b>0.802</b>	<b>0.947</b>	<b>0.942</b>	<b>0.776</b>	<b>0.936</b>	<b>0.930</b>
Baseline Treatment 2		300	<b>0.500</b>	<b>0.933</b>	<b>1.000</b>	<b>0.770</b>	<b>0.941</b>	<b>0.938</b>	<b>0.743</b>	<b>0.940</b>	<b>0.942</b>

**Note.** For each treatment (in rows), the second column gives the number of observations – the number of individual subjects is  $N/10$ . For the first round of play, Rounds 2-10 and overall averages, the three sub-columns provide the average frequency of: reliance from player A (decision  $R$ ), reliability from player B (decision  $r$ ) and reliability conditional on being relied on ( $r|R$ ).

### Aversion to inequality under improved saliency

Table 1.5 summarizes the outcomes from our four companion treatments, along with a reminder of the main results from BT1 and ET1. For the three egalitarian treatments (ET2, ET3 and ET4), we split the results according to the location of the experiment. This allows us to assess the level of cultural specificity in the behavior observed in this game.

In the light of the data from these three treatments, it is difficult to argue that there is any systematic difference between the patterns of behavior in each country. Player As' behavior is only subject to minor discrepancies and seems quite similar in France and Poland. For treatments ET 2 and ET 4, players Bs' behavior in both countries is practically identical. The sole visible difference concerns player Bs' behavior in treatment ET 3, where Polish subjects are slightly more likely to select action  $r$  than French subjects. However, statistical tests are unable to reject the null of equal means in the two countries.<sup>15</sup> We therefore do not consider the cultural background

BT1, and 47.7% in ET1 ( $p=0.249$ ), are coordinated. Cooperative outcomes account for 41.3% and 33% ( $p=0.497$ ), respectively. Type 1 errors are extremely widespread, attaining 39.3% and 39.7% ( $p=0.955$ ) of global outcomes. The most costly Type 2 errors are also pronounced, reaching 7.7% and 12.7% ( $p=0.220$ ) in BT1 and ET1, respectively.

<sup>15</sup>The Kolmogorov-Smirnov test using all session averages does not detect differences between the two countries either in the population of player As ( $p=0.980$ ) or in the population of player Bs ( $p=0.317$ ). We can only test the nullity of the difference for each treatment separately in the first round, when individual decisions are not correlated. The  $p$ -values are  $p=0.569$  in ET2, and  $p=1$  in ET3 and ET4 for player As, and  $p=0.697$ ,  $p=0.695$ ,  $p=0.372$  for

as an influential factor in our experimental data, and accordingly pool both locations in the data analysis below.

In ET3 and ET4, we improve the saliency of the action leading to the Pareto-efficient outcome for each player in turn. As compared to ET1, ET3 increases player Bs' conditional surplus from playing  $r$ , while ET4 also increases player As' conditional surplus from playing  $R$ . Comparing these three experimental conditions, we observe very little variation in player Bs' actions: although reliable decisions are observed slightly more often when incentives become more salient – for instance, in 63% of round 1 decisions in ET1 compared with 80% in ET3 and 77% in ET4 – we cannot identify any statistically significant difference.<sup>16</sup>

Setting ET1 as a benchmark, conditions ET3 and ET4 also allow us to investigate two important issues concerning player As' decision-making. First, do player As react to partners' enhanced incentives to act reliably in ET3 – notwithstanding player Bs' actual neutrality? Second, do player As react to the enhancement of their own incentives to rely on their partners in ET4? Based on our data, the answer to the first question is clearly negative: the improvement of the saliency of player Bs' decisions between ET1 and ET3 only results in a marginal and statistically insignificant drop in the frequency of player As' secure choices.<sup>17</sup> The answer to the second question, on the contrary, appears to be positive. Although ET4 brings no significant improvement in term of player As' reliance in the initial round, the increase in the rate of reliability becomes highly significant in subsequent rounds.<sup>18</sup>

ET2 is designed to detect another potential motivation underlying reliance: as compared with ET4, payoffs generated by a secure choice are equalized, so that we should observe a fall in reliance if player As use decision  $R$  to move away from outcomes with highly unequal payoffs. We clearly reject this hypothesis both in the first round ( $p=0.487$ ) and in subsequent repetitions ( $p=0.539$ ). On the other hand, we find a strong effect on player Bs' behavior after the first round (where the rate of decisions  $r$  equals 84%). In this version of Egalitarian Treatment, subjects in the role of player B attain the highest proportion of decisions  $r$ , with almost 95% of actions  $r$  in rounds 2-10.<sup>19</sup> We complete our design by considering a variation of ET2, in which the payoffs associated

---

player Bs. The results from parametric regressions (see Section 1.5) confirm this conclusion.

<sup>16</sup>The  $p$ -values from mean differences between treatments ET1 and ET3/ET4 are  $p=0.175/p=0.399$  in round 1,  $p=0.491/p=0.461$  in rounds 2-10, and  $p=0.446/p=0.418$  in rounds 1-10. The null hypothesis that behavior is the same in all three treatments cannot be rejected in round 1 ( $p=0.290$ ), in rounds 2-10 ( $p=0.753$ ), or in rounds 1-10 ( $p=0.710$ ).

<sup>17</sup> $p=0.469$  in round 1,  $p=0.249$  in rounds 2-10,  $p=0.193$  in rounds 1-10.

<sup>18</sup>In round 1:  $p=0.469$  against ET1,  $p=0.809$  against ET3; in rounds 2-10:  $p=0.013$  against ET1,  $p=0.004$  against ET3; for rounds 1-10:  $p=0.008$  against ET1,  $p=0.007$  against ET4. Although we cannot reject the null hypothesis that subjects' behavior in round 1 is the same in all three treatments ( $p=0.462$ ), we can do so for rounds 2-10 ( $p=0.006$ ) and rounds 1-10 ( $p=0.007$ ).

<sup>19</sup>Comparing ET2 against ET3:  $p=0.782$  in round 1,  $p=0.081$  in rounds 2-10, and  $p=0.104$  in rounds 1-10. Comparing ET2 against ET4:  $p=0.555$  in round 1,  $p=0.006$  in rounds 2-10, and  $p=0.001$  in rounds 1-10. Although we cannot reject the null hypothesis that subjects' behavior in round 1 is the same in all three treatments ( $p=0.701$ ),

Table 1.6: Distribution of players according to their actions

	N	Number of times player B plays $l$												Number of times player A plays $L$											
		0	1	2	3	4	5	6	7	8	9	10	0	1	2	3	4	5	6	7	8	9	10		
BT1	30	12	4	5	1	3	2	2	1	0	0	0	3	4	4	1	1	4	2	2	1	0	8		
ET1	30	11	4	2	1	5	2	0	3	0	0	2	6	1	2	3	0	4	1	0	1	5	7		
ET3	40	22	5	1	3	4	1	2	0	0	0	2	15	2	4	1	1	1	1	2	0	2	11		
ET4	30	17	2	0	4	3	2	0	0	1	1	0	17	4	0	1	0	0	0	0	2	2	4		
BT2	30	23	0	3	4	0	0	0	0	0	0	0	11	4	2	4	2	2	2	1	0	0	2		
ET2	50	37	4	5	1	0	3	0	0	0	0	0	23	10	3	2	2	3	0	0	1	1	5		

**Note.** For each treatment (in rows) the table reports the distribution of player Bs (player As) according to the number of times decision  $l$  ( $L$ ) is chosen over the 10 repetitions of the game: the cells refer to the number of individual subjects.

with the Pareto-dominant outcome are unequal between players. The comparison between BT2 and ET2 thus complements BT1-ET1 in a context in which payoff differences are salient and the payoff differences are equalized between players in both Nash equilibria of the game. Once again, we find no important inter-treatment differences in either player As' or player Bs' actions.<sup>20</sup>

To get further insight into individual behavior, Table 1.6 groups together player Bs and player As in each treatment according to the number of times they choose (respectively) the weakly-dominated decision  $l$  or the secure and unreliant action  $L$  over the 10 repetitions of the game. The general conclusion we can draw from this distribution is that aggregate decisions are not simply a matter of a few individual outliers. On the side of player Bs' behavior, the treatments can be split into two sub-groups. In BT2 and ET2, on the one hand, 75%-77% of subjects never use the weakly-dominated strategy  $l$ , and the individual frequencies of actions  $l$  amongst the remaining subjects never exceed 50%. In BT1, ET1, ET3 and ET4, on the other hand, the proportion of perfectly reliable player Bs varies between 37% and 57%, and the number of actions  $l$  among remaining subjects is much more dispersed, often exceeding 5/10. As for player As, we can also observe two distinct groups of treatments. The first sub-group, which comprises treatments BT2, ET2, and ET4, is characterized by a large proportion of subjects (between 37% and 57%) who always rely on their partners, and a relatively small proportion of those who never do so (between 7% and 13%). In BT1, ET1 and ET3, by contrast, most observations are dispersed far from the absolute reliance category, and this is coupled with a high share of absolutely unreliant subjects.

Table 1.7: Parametric regressions on the determinants of cooperative behavior

	Pr(R)				Pr(r)			
	Coef.	p-val.	ME	p-val.	Coef.	p-val.	ME	p-val.
<i>Intercept</i>	-1.182	0.108	—	—	0.228	0.565	—	—
<i>Egalitarian Treatment</i>	0.033	0.849	0.006	0.847	-0.058	0.815	-0.010	0.814
Improved saliency for player A	0.348	0.031	0.068	0.020	0.099	0.706	0.017	0.704
— for player B	-0.027	0.883	-0.005	0.881	0.077	0.793	0.013	0.791
<i>Poland</i>	0.081	0.641	0.016	0.635	0.185	0.368	0.032	0.365
<i>Player B Disadvantaged</i>	-0.107	0.598	-0.021	0.591	-0.427	0.071	-0.074	0.051
Learning: player A								
$\mathbf{1}[R_{t-1}]$	2.172	0.000	0.423	0.000	—	—	—	—
Population $B_{t-1}$	0.442	0.579	0.086	0.577	—	—	—	—
Learning: player B								
$\mathbf{1}[r_{t-1}]$	—	—	—	—	1.414	0.000	0.245	0.000
Population $A_{t-1}$	—	—	—	—	-0.164	0.841	-0.028	0.839
<b>Round Dummies</b>								
$t = 3$	-0.025	0.914	-0.005	0.913	0.044	0.733	0.008	0.732
$t = 4$	-0.130	0.493	-0.025	0.487	0.064	0.647	0.012	0.648
$t = 5$	-0.153	0.268	-0.029	0.251	0.012	0.947	0.002	0.946
$t = 6$	-0.145	0.418	-0.028	0.417	0.046	0.757	0.009	0.755
$t = 7$	-0.076	0.542	-0.014	0.533	0.155	0.538	0.027	0.529
$t = 8$	-0.274	0.075	-0.053	0.052	0.365	0.074	0.058	0.067
$t = 9$	-0.098	0.503	-0.019	0.498	-0.010	0.952	-0.002	0.951
$t = 10$	-0.200	0.240	-0.038	0.235	0.166	0.393	0.029	0.377
Log Likelihood	-672.288				-599.371			
Pseudo $R^2$	0.448				0.227			
Prob $>\chi^2$	0.000				0.000			

**Note.** Coefficients and average marginal effects (along with their respective  $p$ -values) from probit regressions on individual decisions. The two models are estimated separately on observations from rounds 2-10 with session-level clustered data and jackknife standard errors. Treatment variables are: *Poland* (=1 for sessions run in Warsaw), *Player B disadvantaged* (=1 for BT2 and ET2), *Improved saliency for player A* (=1 for BT2, ET2, and ET4), *Improved saliency for player B* (=1 for BT2, ET2, ET3, and ET4), *Egalitarian Treatment* (=1 for ET1, ET2, ET3, and ET4).  $\mathbf{1}[R_{t-1}]$  and  $\mathbf{1}[r_{t-1}]$  are dummy variables, =1 if player A (B) played  $R$  ( $r$ ) in the previous period. The variable *Population  $A_{t-1}$*  ( *$B_{t-1}$* ) measures the average rate of reliance (reliability) in all rounds of a given session up to the current one. The number of observations is  $N = 1890$  in both models.

## Robustness analysis

To assess the robustness of the conclusions drawn from the above two-by-two mean comparisons, we now turn to a parametric analysis on individual data pooled over all treatments. The two

we can do so for rounds 2-10 ( $p=0.012$ ) and rounds 1-10 ( $p=0.003$ ).

<sup>20</sup>For player As, we find  $p=0.819$  in round 1 and  $p=0.745$  in rounds 2-10;  $p=0.741$  in rounds 1-10. For player Bs,  $p=0.306$  in round 1 and  $p=0.858$  in rounds 2-10;  $p=0.902$  in rounds 1-10.

outcomes of interest are the two decisions leading to the Pareto-dominant outcome:  $R$  from player A and  $r$  from player B. Due to the nature of these two variables, we run probit regression models. In order to better identify the channels influencing behavior in each treatment, we include treatment variables measuring the specificity of the payoff structure. The indicator *Egalitarian Treatment* is set to 1 when payoffs generated by the dominant outcome are equalized between players (*i.e.*, in all ET treatments; it is 0 for BT treatments). Two dummy variables are used to measure the *improved saliency* of each decision separately for player A – set to 1 when the gap in player As’ payoff between  $(R, r)$  and  $(L)$  amounts to 1.50 euros (treatments BT2, ET2, and ET4) and 0 if it equals only 0.25 euros (BT1, ET1, and ET3) – and for player B – set to 1 in treatments BT2, ET2, ET3 and ET4. The variable *Player B disadvantaged* is set to 1 for treatments BT1, ET1, ET3, and ET4 (to 0 otherwise). Last, we capture potential cultural differences by including an indicator of whether the session took place in Poland (1) or France (0).

Parametric models allow us to take into account learning over time in a more disaggregated way. For each player, we distinguish two possible channels of time dependency. The first is one’s own willingness to choose the decision leading to the Pareto-dominant outcome – measured respectively by the decision to play  $R$  in the previous round for player A,  $\mathbf{1}[R_{t-1}]$ , and the similarly-defined variable  $\mathbf{1}[r_{t-1}]$  for player B. The second is the observation of past decisions in the population of partners – which can be inferred in some instances from the actual payoff earned in the round, which subjects learn at the end of each period. To take this dimension into account, we include the frequency of reliant/reliable decisions taken in a given session in all rounds up to the current one – variables Population  $A_{t-1}$  and Population  $B_{t-1}$ . Because this is conditional on past behavior, we use data from rounds 2 to 10 only. Last, we also take simple time effects into account through round-specific dummies.

The results from separate estimations for each outcome variable are provided in Table 1.7. The main additional insights from the regressions come from the learning variables. Subjects in both roles are conservative in their decision-making: both past reliance and reliability increase the odds of similar behavior in the future. Apart from this effect, none of the time-dependent variables appear to have much influence on behavior. In particular, subjects do not seem to update their decision-making according to the actual behavior in the population of their partner.

The estimates on treatment variables confirm that our conclusions are robust to this conditioning. We can observe no significant effect of whether the decisions are elicited in France or in Poland on either group of players. For player As, the only influential dimension of the payoff structure is the saliency of the opportunity cost of the secure choice (which is improved in treatments BT2 and ET2-ET4). For player Bs, saliency does not have any significant effect. The only significant change in behavior is obtained once the strategic disadvantage induced by low payoffs in the dominated outcome is eliminated, as in treatments BT2 and ET2. More importantly for our main purpose, equalizing payoffs in the Pareto-dominant outcome does not change either player

Bs' or player As' behavior.

## Discussion

To sum up, the empirical answers to our main research questions are clearly negative: our data unambiguously reject the inequality aversion hypothesis, previously proposed as a potential driving force behind the two puzzling behaviors we study. Through the replication of some of our treatments in an East European country, we also confirm that cultural differences – due, for example, to the focus on WEIRD (Western, Educated, Industrialized, Rich, and Democratic) people, as defined by Henrich, Heine, and Norenzayan (2010) – do not drive the observed behavior. This result is in line with the robustness experiments performed by Beard, Beil, and Mataga (2001) on Japanese subjects. Again in accordance with previous evidence (as summarized in Section 1.1), we also find that the saliency of reliance is an important driving force of the willingness to rely on partners. Last, all these results are robust to repetition-based learning through several periods of play of the one-shot game.

The only instance in which we observe a strong effect of the treatment on behavior is when the equality of payoffs is the same in both Nash equilibria. This also affects the default earning of player Bs when player As decide alone on the final outcome (by selecting the secure choice) – which can be seen as a natural reference point for player Bs. Observed behavior in these last two treatments thus echoes the literature on money-burning behavior, showing that people sacrifice their own wealth to reduce the wealth of others who have benefited from an exogenous procedural advantage (see, *e.g.*, Zizzo and Oswald, 2001).<sup>21</sup> According to this interpretation, what we observe could thus be the result of player Bs' willingness to punish player As for procedural injustice.

Still, we find that the rate of secure choices by player As does not vary as a function of player Bs' behavior, and remains high even in settings where player Bs commonly use the payoff-maximizing strategy. This means that the fear that player Bs' behavior is driven by social preferences (taking the form of either the aversion to inequality or the propensity to burn other's money) is definitely not the reason why suboptimal outcomes are observed in this class of games. Our data thus restrict the set of available hypotheses to only a few candidates. The first explanation is that the behavior of player As is induced by strategic uncertainty, which goes along the lines of Rosenthal's genuine conjecture. This path is explored in the remainder of this chapter. The second avenue is that player As react non-strategically to the decision problem, so that their choices are unaffected by player Bs' behavior. Such non-strategic play may occur due to, among others, the failure of game form recognition (Chou, McConnell, Nagel, and Plott, 2009); misconceptions (Plott and Zeiler, 2005); subjects' confusion about the strategic context (Ferraro and Vossler, 2010); or, last, by a lack of commitment from subjects towards the experiment they are involved in and their

---

<sup>21</sup>In accordance with our results on the effect of saliency, Zizzo and Oswald (2001) also show that the price elasticity of the demand for such punishment is very low.

Table 1.8: The experimental game

Player A	Player B	
	$l$	$r$
$L$	(9.75;3)	(9.75;3)
$R$	(3;4.75)	(10;5)

unwillingness to take the decision problem seriously enough (see, *e.g.*, Jacquemet, Joule, Luchini, and Shogren, 2009, 2012). This direction is examined in Chapter 2.<sup>22</sup>

### 1.3 Coordination under enhanced information

May inefficient behavior in Rosenthal’s coordination game be contained by enhancing the amount of information that is held by players about their interaction partners? In order to address this question, our experimental treatments vary what players know about their interaction partners, while holding constant the payoff structure of the game. The first kind of information is a simple repetition of the game. Pairs are rematched according to a perfect stranger procedure, which, over time, enhances players’ knowledge about the general population of their partners. Two further treatments implement specific information about the current interaction partner. One treatment allows second movers to provide first movers with “soft” information about their intended behavior through cheap-talk communication: in this communication treatment, second movers send closed-form messages to first movers prior to the decision-making stage of the game. In the observation treatment, we turn to “harder” information in the form of actual decisions taken by second movers: before taking their own decisions, first movers are informed about all the decisions their current interaction partners took in the past. Importantly, previous experimental studies of communication and observation in experimental coordination tasks commonly suggest that such institutions help overcome coordination failures and foster efficient cooperation.<sup>23</sup>

#### 1.3.1 Analysis of the game

For the sake of replication, our core game is based on the original experiment of Beard and Beil (1994) and furthermore relies on the design introduced in the previous section. Among the various

<sup>22</sup>Note that the two groups of explanations lead to very different conclusions: the first one raises a caveat of game theory’s ability to describe behavior, while the other challenges the internal validity of experiments applied to this class of games.

<sup>23</sup>See, *e.g.*, Cooper, DeJong, Forsythe, and Ross (1992), Crawford (1998), Charness (2000), Duffy and Feltovich (2002), Duffy and Feltovich (2006), Blume and Ortmann (2007).



payoff combinations we selected matrix BT1 (presented in Table 2.1) which coincides with Beard and Beil’s treatment 1 and exhibits several attractive features: *(i)* as in the original setting, it does not lead to any conflict of interests between partners; *(ii)* the rate of player As’ unreliable choices related to this payoff structure is remarkable: 65.7% in Beard and Beil’s experiment, 51% in the experiment presented in Section 2; and *(iii)* this is the only payoff structure where important deviations from the dominant strategy by player Bs were reported by both the genuine study by Beard and Beil (in 17% of all cases where player A made a reliant decision  $R$ ) and our earlier experimental investigation (attaining 20% of all actions taken by player Bs in BT1).

Before moving to the details of our experimental treatments, let us briefly recapitulate the main properties of the game presented in Table 2.1. In this one-shot, simultaneous-move game one player has a weakly dominated strategy: for player B, playing  $l$  is never payoff improving. Because the dominance is only weak, all outcomes are rationalizable: for each action, there is some belief about opponent’s behavior to which this action is a best reply. Only two of them are pure strategy Nash equilibria:  $(R, r)$ , which is payoff dominant; and  $(L, l)$ , which arises because  $L$  is the appropriate action against a player B choosing  $l$ . In fact,  $L$  is a risk-free action for player A; and the strategic risk associated with playing  $R$  is quite extreme: while payoffs slightly increase should the most efficient outcome be reached when  $R$  is played, they strongly decrease as compared to the safe choice  $L$  in the case of a coordination failure  $(R, l)$ . As stressed above, existing lab implementations of this game observed a high rate of unreliance (increasing in the difference in payoff between the two equilibria) and little unreliability conditional on being relied on (decreasing in the difference in payoff between the two actions).

## Learning

In the baseline treatment we allow both players to experiment with the features of the game through trials and errors across several repetitions of the one-shot game (see Section 1.3.2 for a detailed description of the design). Because the payoffs are asymmetric, the room for learning is very different for both players. On player Bs’ side, one could think that the use of weakly dominated strategies mainly comes from inexperienced players. If it is true, we should observe a decrease in the share of decisions  $l$  as players gain some experience. For player As, by contrast, the choice between  $L$  and  $R$  depends on the anticipated behavior of their current interaction partners. As player As meet different player Bs from one repetition to the other, the baseline allows for learning the actual distribution of behaviors in the population of player Bs.

## Words

In the communication treatment, we introduce specific information about the current interaction partner through one-way, closed-form communication. The effect of communication on the outcomes of the game depends on the link between messages and future actions. To that regard, the

theoretical properties of cheap-talk from player B to player A in this game are an illustration of the Aumann (1990) critique towards Farrell (1988). Using the Farrell and Rabin (1996) nomenclature, the message “I will choose  $r$ ” is self-committing because if it is enough to convince player A to play  $R$ , it subsequently creates the incentives for B to play in accordance with the message. But the message is not self-signalling: player B always wants player A to choose  $R$ , whether his true intention is to play  $r$  or not. A rational player A should thus anticipate that the message announcing action  $r$  is only meant to lead him to choose  $R$  and does not tell anything about the true intentions of player B. As a result, communication should not matter in this game.

Ellingsen and Östling (2010) study cheap-talk in such contexts by relying on non-rationalizable beliefs. They formalize the idea originally introduced by Crawford (1998) that communication provides reassurance to the receiver about the degree of rationality of the sender. Ellingsen and Östling (2010) relax the assumption that players believe with probability 1 that their opponents are rational. The maintained assumption is that players have a weak (lexicographic) preference for being honest – *i.e.* players are truthful whenever they are indifferent between messages. In that case, all players signal their true intentions. In our game, this implies the message “I will choose  $r$ ” can be used by player Bs to signal their ability to disregard weakly dominated strategies, hence reassuring player As about the actual risk induced by seeking for the payoff dominant outcome.

In the experimental literature devoted to the effect of cheap-talk communication, many papers focus on the Stag Hunt game, which shares some important features with the Rosenthal game studied here: there is a conflict between efficiency and strategic risk, and (in the most standard version of the game) one-way communication is self-committing but not self-signalling. In this context, one-way communication between players is generally found to foster efficiency to a very large extent – as compared to the outcomes without communication.<sup>24</sup> Based on the theoretical properties discussed above, as well as this previous evidence, we thus expect cheap-talk to improve the outcomes of the game, through a reassurance effect that increases the completeness of information between players.

## Actions

Our third experimental condition follows a recent experimental literature contrasting the coordination properties of communication with the performance of observation of current interaction partner’s past actions. The most important difference with the information coming from cheap-talk communication is that past actions are not cheap: the signal is costly because it comes from actual decisions associated with variations in payoffs. Duffy and Feltovich (2002) introduce both cheap-talk and observation of partner’s most recent action in three 2 x 2 games: Prisoner’s

---

<sup>24</sup>See Crawford (1998) for an earlier survey of the theoretical and experimental literature on cheap-talk, Ellingsen and Östling (2010) for a detailed survey of the evidence in Stag Hunt games, and Cooper, DeJong, Forsythe, and Ross (1992) and Charness (2000) for a comprehensive experimental study.

Dilemma, Stag Hunt and Chicken.<sup>25</sup> Stag Hunt is the closest to our setting. In the version of the Stag Hunt game they implement, the safe action leads to the same payoff regardless of opponent's choice, so that messages are highly credible (*i.e.* both self-signalling and self-committing). It is not the case in the Chicken game, where messages are self-committing but not self-signalling (like in our game), but the nature of strategic interaction is different from the one we study here, because there is no Pareto dominant outcome. Duffy and Feltovich (2002) show both treatments lead to an increase in frequency of the Nash equilibria and improve the efficiency of outcomes whatever the structure of the game. Cheap-talk appears to be more effective than observation in the Stag Hunt game – “words speak louder than actions”. Observation brings about better results in the Chicken game – “actions speak louder than words”. The game we study is different from both of them: as in Stag Hunt, one equilibrium is payoff dominant, but as in Chicken, communication is not self-signalling. How observation and communication will perform to improve the outcomes in our game is thus still an open question. Our observation treatment moreover increases the amount of information as compared to previous studies: we provide players with the full history of past decisions, rather than only the previous action. The design of this treatment hence enables us to identify the weight of each historical component in the overall perception of reputation.

### 1.3.2 Experimental design

The Baseline condition implements a simple repetition of the one-shot game presented in Table 2.1 over an undefined number of rounds. We study the role of information in the game through two information boosting devices applied prior to the decision-making stage of every interaction: cheap-talk messages from Bs to As and historical information on how each player B acted in all the previous periods.

#### Baseline treatment

Our focus on the effect of information in the game led us to introduce several modifications to the original experiment: *(i)* the one-shot game is played repeatedly by experimental subjects, and *(ii)* we implement the normal form of the game rather than the genuine sequential form.

The experiment involves 10 rounds, each consisting of the core game presented in Table 2.1. Roles are fixed, so that each subject takes 10 decisions as either player A or player B. The experiment was designed so as to remain as close as possible to a one-shot game. First, the pairs are rematched after each round using a round-robin perfect stranger design (each session involves 20

---

<sup>25</sup>There is still few papers that implement this kind of feedback information. Bracht and Feltovich (2009) apply these two treatments in the Gift-Exchange Game. The results show a striking contrast between treatments: while observation is effective in reinforcing cooperation, the effect of communication visibly lags behind. For further discussion of the effects of communication and observation on strategic behavior in the lab, see also Çelen, Kariv, and Schotter (2010).

subjects). Second, although the number of repetitions is pre-determined, we do not reveal it to subjects in order to avoid end-game effects – in the experimental instructions we only inform that the game contains several rounds. Last, we associate take-home earnings from the experiment with only one round out of ten. For that matter, one round is randomly drawn at the end of the experiment (the same for all subjects).

To ensure the homogeneity of rounds despite repetition, we also modify the sequentiality of the game originally introduced by Rosenthal (1981). As pointed out by Binmore, McCarthy, Ponti, Samuelson, and Shaked (2002), the repetition of one-shot multi-stage games may induce some unwarranted heterogeneity and selection bias in observed behavior, because players are induced to distinguish between rounds based on the decisions made in earlier stages of the game. Unlike the original Beard and Beil (1994) experiment, we thus ask both player A and player B to make decision in each period. To make it as close as possible to the original sequential game, we describe the decision phase to subjects as follows:<sup>26</sup> player A is first asked to choose between  $L$  and  $R$ , then player B chooses between  $l$  and  $r$ . Payoffs depend only on player A’s decision should  $L$  has been chosen, or on both players’s decision otherwise.

To sum up, our Baseline treatment implements a repeated version of the one-shot game originally studied by Beard and Beil (1994). We use a perfect stranger design, the normal form of the game, an unknown termination rule and a one-round compensation rule to avoid that subjects compute the expected value of the entire game. This should induce players to maximize their utility in each repetition of the one-shot game. As a first step towards assessing the role of information in the Rosenthal puzzle, intra-comparisons in the Baseline treatment hence allow us to assess the robustness of the results to repetition. We further increase the amount of information in two subsequent experimental treatments.

## Experimental treatments

**Pre-play communication.** This treatment allows player Bs to provide information about their intended play to player As. In every round, prior to the decision-making phase, player B has to send a message to player A. In the experimental implementation of cheap-talk messages, several trade-off must be solved. As raised by, *e.g.*, Farrell and Rabin (1996), cheap-talk ought to be meaningful, *i.e.* to have a precise meaning. Messages of the “I will do ...”-type, which might be considered as a bit oversimplified, are nonetheless highly meaningful. Voluntary free-form communication, by contrast, improves the informational content of communication but always gives the sender an opportunity to send an empty message or a message that is either meaningless or imprecise – which is hard to interpret for both the receiver and the experimenter.<sup>27</sup> Given that our primary

---

<sup>26</sup> An English translation of the original instructions in French is provided as supplementary material in Section 1.5.

<sup>27</sup> Experimental results from Charness and Dufwenberg (2006, 2008) substantiate that impersonal messages, that have been prefabricated by the experimenter, work effectively in coordination games, whereas in trust games a more customized free-form communication seems to be needed. Similarly, Bochet, Page, and Putterman (2006) find that

goal is to boost the amount of information, we want to encourage players to communicate in a precise and clear manner. We hence implement a fixed-form communication and limit the set of possible messages to three options only, out of which two contain precise information, while the third is empty. Before any decision takes place in the round, player B is asked to choose one out of the three following messages:

☐ I will choose  $r$                       ☐ I will choose  $l$                       ☐ I will choose either  $l$  or  $r$

by clicking on the relevant field on her computer screen. This message is then displayed on player A's computer screen. Once player A confirmed the reception of the message, the round moves to the decision phase. It is highlighted in the written instructions that messages are not binding (decisions from player Bs can be anything following any of the messages) and do not affect experimental earnings.

**Observation of historical information.** In the third condition, we allow subjects in the role of player As to inspect all the decisions made by their current interaction partner in all previous rounds. In every round, before the decision-making phase, player B is asked to wait while player A is provided with the history of choices made by player B. Following, *e.g.*, Bolton, Katok, and Ockenfels (2004) we make available the full history of past decisions rather than only the last one (see, *e.g.*, Bracht and Feltovich (2009), Duffy and Feltovich (2002)). In each round, player As thus receive a list of all the decisions made so far by their current interaction partner. Since pairs are rematched before each round, this information is updated and extended accordingly. Once player A confirmed to be aware of player B's history, the decision-making phase starts.

## Experimental procedures

For each treatment, we ran three sessions involving 20 subjects each. Upon arrival, participants are randomly assigned to their computers and asked to fill in a small personal questionnaire containing basic questions about their age, gender, education, etc. The written instructions are then read aloud. Players are informed that they will play some (unrevealed) number of rounds of the same game, each round with a different partner, and that their own role will not change during the experiment. Before starting, subjects are asked to fill in a quiz assessing their understanding of the game they are about to play. Once the quiz and all remaining questions are answered, the experiment begins. Prior to the first round, players are randomly assigned to their roles – either A or B. They are then anonymously and randomly matched to a partner and asked for their choice,  $R$  or  $L$  for player As, and  $r$  or  $l$  for player Bs. At the end of each round, each subject is informed solely about her own payoff. Once all pairs complete a round of the game, subjects are informed whether a new round starts. In this case, pairs are rematched according to a perfect stranger

---

free-form communication yields higher efficiency in a VCM game than numerical messages.

Table 1.9: Summary of experimental evidence on related games

Experiment	Payoff			Observed outcomes				Nb.	
	( $L$ )	( $R, r$ )	( $R, l$ )	$L$	( $R, r$ )	( $R, l$ )	( $r R$ )	r	obs.
Beard and Beil (1994)-Tr.1	(9.75; 3)	(10; 5)	(3; 4.75)	66%	29%	6%	83%	—	35
Beard et al. (2001)-Tr.1	(1450; 450)	(1500; 750)	(450; 700)	79%	18%	3%	86%	—	34
Goeree and Holt (2001)-Tr.2	(80; 50)	(90; 70)	(20; 68)	52%	36%	12%	75%	—	25
Goeree and Holt (2001)-Tr.3	(400; 250)	(450; 350)	(100; 348)	80%	16%	4%	80%	—	25
Baseline, round 1	(9.75; 3)	(10; 5)	(3; 4.75)	77%	23%	0%	100%	80%	30
Baseline, rounds 2-10	(9.75; 3)	(10; 5)	(3; 4.75)	48%	43%	9%	84%	81%	270
Baseline, overall	(9.75; 3)	(10; 5)	(3; 4.75)	51%	41%	8%	84%	81%	300

**Note.** The monetary payoffs displayed in the first three columns are in USD in Beard & Beil (1994), in cents of USD in Goeree & Holt (2001), in Yens in Beard et al. (2001) and in Euros in our treatments.

round-robin matching procedure (in which any pair meets only once in the session). At the end of the experiment, one round is randomly drawn and each player receives the amount in Euros corresponding to her gains in that round, plus a show-up fee equal to 5 Euros.

All sessions took place in the lab of University Paris 1 Panthéon-Sorbonne (LEEP) in between June 2009 and March 2010. The recruitment of subjects has been carried out via LEEP database among individuals who have successfully completed the registration process on Laboratory’s web-site.<sup>28</sup> The experiment involved a total group of 180 subjects, 90 males and 90 females.<sup>29</sup> 86% of them were students, among which 85 subjects are likely to have some background in game theory due to their field of study.<sup>30</sup> 36% never took part in any economic experiment in LEEP before. Participants’ average age is roughly 24. No subject participated in more than one experimental session. Each session lasted about 45 minutes, with an average payoff of 12€ (including a 5€ show-up fee).

### 1.3.3 Results

The last three rows of Table 2.2 provide a summary of observed behavior in our baseline treatment along with results from previous experimental studies using similar payoff structures (top part of the Table). Our results are generally in line with what has been observed in other studies, despite the differences in the design described in Section 1.3.2: the average rate of unreliance is 51%, very close to the one observed in Goeree and Holt (2001) who apply the strategy method to the

<sup>28</sup>The recruitment uses ORSEE (Greiner, 2004); the experiment is computerized through a software developed under REGATE (Zeiliger, 2000).

<sup>29</sup>This 50-50 spread of genders is purely incidental.

<sup>30</sup>Disciplines such as economics, engineering, management, political science, psychology, mathematics applied in social science, mathematics, computer science, sociology.

Table 1.10: Overall effects of information treatments

		Decisions		Outcomes		Errors	
		Reliant A ( $R$ )	Reliable B ( $r$ )	Cooperation ( $R, r$ )	Coordination ( $R, r$ ) $\cup$ ( $L, l$ )	Type I ( $L, r$ )	Type II ( $R, l$ )
Baseline	1	0.233	0.800	0.233	0.433	0.567	0.000
	2-4	0.456	0.844	0.411	0.522	0.433	0.044
	5-7	0.589	0.767	0.489	0.622	0.278	0.100
	8-10	0.511	0.811	0.400	0.478	0.411	0.111
	Average	0.490	0.807	0.413	0.530	0.393	0.077
Communication	1	0.500	0.800	0.433	0.567	0.367	0.067
	2-4	0.467	0.811	0.422	0.567	0.389	0.044
	5-7	0.678	0.778	0.589	0.722	0.189	0.089
	8-10	0.667	0.811	0.600	0.722	0.211	0.067
	Average	0.593	0.800	0.527	0.660	0.273	0.067
Observation	1	0.167	0.767	0.033	0.133	0.733	0.133
	2-4	0.544	0.833	0.511	0.644	0.322	0.033
	5-7	0.578	0.800	0.511	0.644	0.289	0.067
	8-10	0.589	0.844	0.544	0.656	0.300	0.044
	Average	0.530	0.820	0.473	0.597	0.347	0.057

**Note.** For each treatment, we separate data according to the stage of the game – round 1, rounds 2-4, rounds 5-7, rounds 8-10, and the overall results (consecutive rows). In each row, we provide empirical frequencies of decision  $R$  by player A, decision  $r$  by player B, cooperation on the most efficient Nash equilibrium ( $R, r$ ), coordination on the existing Nash equilibria ( $R, r$ ) or ( $L, l$ ), Type I error ( $L, r$ ) and Type II error ( $R, l$ ) (columns 1-6, respectively).

sequential game. Unsurprisingly, the one-shot sequential games of Beard and Beil (1994); Beard, Beil, and Mataga (2001) are much better replicated by the first round of our baseline treatment than by the overall rate produced by the repetition of the game. Moreover, in all studies the likelihood of observing action  $r$  is alike and equal to roughly 80%. Once all repetitions of our baseline treatment are pooled, the outcomes look consistent with all the previous studies in terms of both efficient coordination – outcome ( $R, r$ ) – and strategic mismatches resulting in outcome ( $R, l$ ). In what follows, we discuss whether and how information helps improve efficiency and overcome coordination failures.

### Aggregate treatment effects

Analogously to the previous section, the statistical inference is based on between-treatment tests for the significance of differences in the proportions of each outcome of interest. We use two kinds of statistical tests. In round 1, where all observations are independent, we use two-sided Fisher's exact tests. In rounds 2-10, individual observations may be correlated within each session due to the re-matching of subjects from one round to the other. We take into account this data structure by using the following parametric procedure. First, we run an OLS regression of

treatment dummies on a dummy dependent variable representing the outcome of interest. In this setup, coefficients correspond to the proportions of outcomes of interest in a given treatment, so that two-sided  $t$ -tests and  $F$ -tests on coefficients allow us to test linear hypotheses about the equality of proportions between treatments. To take into account the within-session correlation, we cluster data at the session level, and use leave-one-out jackknife standard errors to deal with the issue of potential small-sample bias. This method is applied to data from rounds 1-10 and 2-10.

Table 1.10 summarizes aggregate behavior elicited in each of the three treatments. For each treatment, we separate data into five categories related to the stage of the game – the initial round, rounds 2-4, rounds 5-7, rounds 8-10, and finally the overall results (consecutive rows).<sup>31</sup> The first two columns of the table summarize unconditional average behavior of player As and Bs. The right-hand-side describes the resulting outcomes: positive ones (efficiency and coordination) in the two middle columns, and failures in the last two. Thanks to our design, we are able to observe two sources of coordination failure: beyond the outcome arising when player A mistakenly relies on player B, which we classify as type II errors, we also account for type I errors – *i.e.* cases where player A should have relied on player B, since player B would have proved reliable in this case – resulting in outcome  $(L, r)$ .

The likelihood of each outcome depends on both players' behavior. As shown in the second column of Table 1.10, the behavior of player Bs is fairly stable regardless of time and experimental treatments.<sup>32</sup> As a result, any difference between treatments and between rounds is very unlikely to be driven by changes in player Bs' behavior. If anything, the treatment effects of information occur because of changes in the way player As perceive player Bs, rather than through discrepancies between populations of player Bs.

Table 1.11 provides the results of several specifications of the linear probability model described in Section 1.2.2. Model 1 summarizes the changes due to repetition-based learning. Models 2 and 3 provide statistical tests of the effects of treatments. To sum up, the repetition of the one-shot game improves efficiency in the baseline, but at the same time increases the odds of type II errors. Specific information (through communication or observation) reduces the likelihood of coordination failures, but does not improve the efficiency of outcomes. coordination. We comment below on the main driving forces behind these results. We then turn to the way specific information about the interaction partner is accounted for by player As.

---

<sup>31</sup>A round by round description of behavior is provided as supplementary material, Section 1.5.

<sup>32</sup>Fisher's exact test does not reject the null hypothesis that player Bs' decisions in round 1 come from the same distribution in all treatments ( $p=1.000$ ). Model 3 in Table 1.11 suggests that in neither treatment the average proportion of decisions  $r$  in rounds 2-10 significantly differs from the initial round – BT:  $p=0.952$ ; CT:  $p=1.000$ , OT:  $p=0.318$ . Finally, on the basis of Model 1 in Table 1.11 we also test a joint hypothesis that the means in all treatments are statistically different in rounds 2-4 through  $H_0 : (BT\_rounds2 - 4 = CT + CT\_rounds2 - 4) \cap (BT\_rounds2 - 4 = OT + OT\_rounds2 - 4)$ . No difference arises either in this early stage ( $p=0.925$ ), or in rounds 5-7 ( $p=0.932$ ) or rounds 8-10 ( $p=0.917$ ).



Table 1.11: Statistical support to Table 1.10

	Reliant A Pr( $R$ )		Reliable B Pr( $r$ )		Cooperation Pr( $R, r$ )		Coordination Pr( $(R, r) \cup (L, l)$ )		Type I errors Pr( $L, r$ )		Type II errors Pr( $R, l$ )	
	coef	p-val.	coef	p-val.	coef	p-val.	coef	p-val.	coef	p-val.	coef	p-val.
Model 1												
Intercept	0.233	0.000	0.800	0.000	0.233	0.000	0.433	0.003	0.567	0.001	0.000	0.410
BT_rounds2-4	0.222	0.000	0.044	0.820	0.178	0.001	0.089	0.620	-0.133	0.484	0.044	0.009
BT_rounds5-7	0.356	0.000	-0.033	0.717	0.256	0.000	0.189	0.173	-0.289	0.032	0.100	0.002
BT_rounds8-10	0.278	0.001	0.011	0.890	0.167	0.022	0.044	0.764	-0.156	0.266	0.111	0.043
CT	0.267	0.178	0.000	1.000	0.200	0.202	0.133	0.491	-0.200	0.385	0.067	0.122
CT_rounds2-4	-0.033	0.836	0.011	0.880	-0.011	0.941	0.000	1.000	0.022	0.898	-0.022	0.122
CT_rounds5-7	0.178	0.521	-0.022	0.838	0.156	0.563	0.156	0.528	-0.178	0.476	0.022	0.122
CT_rounds8-10	0.167	0.555	0.011	0.816	0.167	0.553	0.156	0.531	-0.156	0.535	0.000	1.000
OT	-0.067	0.256	-0.033	0.816	-0.200	0.006	-0.300	0.025	0.167	0.164	0.133	0.122
OT_rounds2-4	0.378	0.077	0.067	0.290	0.478	0.001	0.511	0.000	-0.411	0.007	-0.100	0.332
OT_rounds5-7	0.411	0.013	0.033	0.412	0.478	0.001	0.511	0.000	-0.444	0.005	-0.067	0.172
OT_rounds8-10	0.422	0.056	0.078	0.308	0.511	0.001	0.522	0.000	-0.433	0.006	-0.089	0.365
Model 2												
Intercept	0.490	0.000	0.807	0.000	0.413	0.000	0.530	0.000	0.393	0.000	0.077	0.000
CT	0.103	0.110	-0.007	0.941	0.113	0.211	0.130	0.018	-0.120	0.010	-0.010	0.779
OT	0.040	0.744	0.013	0.882	0.060	0.659	0.067	0.403	-0.047	0.507	-0.020	0.202
Model 3												
Intercept	0.233	0.000	0.800	0.000	0.233	0.000	0.433	0.003	0.567	0.001	0.000	0.658
BT_rounds2-10	0.285	0.000	0.007	0.952	0.200	0.000	0.107	0.458	-0.193	0.203	0.085	0.000
CT	0.267	0.178	0.000	1.000	0.200	0.202	0.133	0.491	-0.200	0.385	0.067	0.122
CT_rounds2-10	0.104	0.656	0.000	1.000	0.104	0.648	0.104	0.626	-0.104	0.637	0.000	1.000
OT	-0.067	0.256	-0.033	0.816	-0.200	0.006	-0.300	0.025	0.167	0.164	0.133	0.122
OT_rounds2-10	0.404	0.037	0.059	0.318	0.489	0.001	0.515	0.000	-0.430	0.005	-0.085	0.289

**Note.** Columns summarize the results of session-clustered (9 clusters in total, 100 observations per cluster, standard errors corrected with delete-one jackknife) OLS regressions of treatment-related variables on player A's decision  $R$ , player B's decision  $r$ , and outcomes: cooperative  $(R, r)$ , coordinated  $(R, r) \cup (L, l)$ , Type I errors  $(L, r)$  and Type II errors  $(R, l)$  (columns 1-6, respectively). The intercept represents the reference frequency in round 1 of the Baseline treatment; dummies CT and OT correspond to the change in the intercept due to the Communication and observation treatments. Remaining coefficients are interpreted as an absolute change in the frequency of dependent variables due to achieving certain stages of the game. Prefix BT stands for the baseline treatment, CT and OT for the Communication and observation treatments, respectively.  $p$ -values come from two-tailed  $t$ -tests for nullity of coefficients.

We first focus on repetition-based learning by comparing outcomes across rounds within the baseline treatment. As shown in Table 1.10, the rate of reliant decisions from player As more than doubles between round 1 and the subsequent occurrences of the game. Given the stability of player Bs' actual decisions, this suggests that over time player As update their beliefs about the population of player Bs. The main effect of this change in player As' behavior is an important improvement in the share of the efficient outcome, from 23% of first round outcomes to 43% of subsequent repetitions (the difference is significant at the 5% level according to Model 1). This comes at a price in terms of coordination failure: while the risk of type I errors falls (significantly only for rounds 5-7), type II error gets more likely (the rise is significant at the 5% level for all triplets of periods).

In the communication treatment, all outcomes become much less sensitive to the repetition of the game. As compared to the baseline situation, cheap-talk strongly increases the reliance rate in the first round, from 23% in the baseline treatment to half of decisions in the communication treatment ( $p=0.060$  using Fisher's exact test), and only slightly in further repetitions of the game,

from 52% to 60% ( $p=0.269$  after testing  $H_0 : BT\_rounds2 - 10 = CT + CT\_rounds2 - 10$  in Model 3). The overall increase in the share of efficient outcomes as compared to the baseline is not significant ( $p=0.211$ , see Model 2). Note, however, that the proportion of efficient outcomes in the communication treatment in the first round is already very close to the one attained due to repetition in the baseline condition. The main effect of cheap-talk communication is an improvement in coordination which is significant at the 5% level (see Model 2). While the likelihoods of both type I and type II errors decrease, only the former is statistically significant, *i.e.* player As are less likely to mistakenly choose the secure option.

In contrast to communication, the observation treatment does not provide specific information to player As at the beginning of the game – the signal becomes available in round 2. The treatment appears anticipated by player As: the rate of reliance is lower in the first round as compared to baseline treatment, which hinders both cooperation ( $p=0.052$  using Fisher's exact test) and coordination ( $p = 0.020$ ). Once information becomes available – in rounds 2-10 – outcomes improve as compared to the first round of the baseline treatment to reach levels similar to those observed after several repetitions.<sup>33</sup> The only exception is type II error being significantly less frequent than in rounds 2-10 of the baseline ( $p = 0.015$ ), so that player As' risk of relying in vain on a partner falls. As compared to communication treatment, the proportion of actions  $R$  in observation condition falls drastically in the first round ( $p = 0.013$  using Fisher's exact test), which boosts the rate of type I errors ( $p=0.009$ ) and reduces coordination and cooperation (both  $p \leq 0.001$ ). In subsequent rounds, however, both experimental conditions provide similar outcomes.<sup>34</sup>

### Informational content of signals

Table 1.12 reorganizes data according to the flow of information. As a benchmark, the first column of the table summarizes the outcomes observed in the baseline treatment when all rounds are pooled. For the communication treatment, the observations are conditioned on the message received by player A – "*I will choose r*", "*I will choose l*", "*I will choose either l or r*". For the observation treatment, we use the reputation of each player B to separate the population into two groups: highly reliable ones and others. For that matter, we construct a reputation index for each player B equal to the rate of decisions  $r$  amongst all decisions made prior to the current round.

<sup>33</sup>Based on Model 3, we assess the effect of observation against baseline through tests of  $H_0 : BT\_rounds2 - 10 = OT + OT\_rounds2 - 10$ . The differences are insignificant as regards reliance ( $p = 0.709$ ), cooperation ( $p = 0.544$ ) coordination ( $p = 0.237$ ) and type I errors ( $p = 0.421$ )

<sup>34</sup>We use Model 1 to test the joint hypothesis that in every triplet of rounds – 2-4, 5-7 and 8-10 – a given outcome is equally frequent in both treatments, that is  $H_0 : (CT + CT\_rounds2 - 4 = OT + OT\_rounds2 - 4) \cap (CT + CT\_rounds5 - 7 = OT + OT\_rounds5 - 7) \cap (CT + CT\_rounds8 - 10 = OT + OT\_rounds8 - 10)$ . We find  $p = 0.688$  for reliance,  $p = 0.987$  for reliability,  $p = 0.669$  for cooperation,  $p = 0.360$  for coordination,  $p = 0.531$  for type I error, and  $p = 0.949$  for type II error.

Table 1.12: Informational content of signals

	Baseline	Communication			Observation		
		$m(r)$	$m(l)$	$m(l/r)$	$B_P$	$B_{IP}$	Unknown
Frequency within treatment	100%	75.7%	12.0%	12.3%	49.3%	40.7%	10%
Reliant A ( $R$ )	49.0%	72.2%	16.7%	21.6%	77.7%	32.0%	16.7%
Reliable B ( $r$ )	80.7%	90.3%	38.9%	56.8%	93.9%	68.9%	76.7%
Cooperation ( $R, r$ )	41.3%	65.6%	8.3%	16.2%	75.7%	23.8%	3.3%
Coordination ( $(L, l) \cup (R, r)$ )	53.0%	68.7%	61.1%	54.1%	79.7%	46.7%	13.3%
Type I error ( $L, r$ )	39.3%	24.7%	30.6%	40.5%	18.2%	45.1%	73.3%
Type II error ( $R, l$ )	7.7%	6.6%	8.3%	5.4%	2.0%	8.2%	13.3%
Nb of observations	300	227	36	37	148	122	30

**Note.** For each treatment in column, the rows provide the proportion of observed decisions (first two rows) and outcomes (last four rows). The first column pools all observations from the baseline. In the middle columns, data from all rounds of the communication game are split according to the message received by player A: "I will choose  $r$ ", denoted  $m(r)$ , "I will choose  $l$ ",  $m(l)$ , and "I will choose either  $l$  or  $r$ ",  $m(l/r)$ . For the observation game (right-hand side of the table), observations are classified according to the reputation of player B: in rounds 2-10, the reputation is perfect, and denoted  $B_P$ , if all previous decisions are  $r$ ; and imperfect otherwise, denoted  $B_{IP}$ . Reputation is unknown in round 1.

We classify player Bs in each round by comparing their reputation to the cut-off probability of decision  $r$  (0.964) that makes a risk neutral player A indifferent between choosing  $L$  and  $R$ . Hence, prior to entering an interaction, each player B may either have a perfect reputation (denoted  $B_P$ ) or an imperfect record ( $B_{IP}$ ).<sup>35</sup> The last column provides observations from the first round, in which no information is available.

For each treatment involving specific information about partner, this classification can thus organize data according to three kinds of informational content delivered to player As: a positive signal – a reassuring message  $m(r)$  in the communication treatment and a perfect reputation  $B_P$  in the observation treatment, a negative signal – one of the non-reassuring messages  $m(l)$ ,  $m(l/r)$ , and an imperfect reputation  $B_{IP}$ , or lack of information – like in the first round of observation treatment. Table 1.13 summarizes statistical tests for differences of proportions conditional on source and content of information, based on parametric regressions of outcomes on signal-related dummy variables discussed above.

In the communication treatment, the frequencies of empty messages and messages announcing the weakly dominated action are both equal to around 12%. 90% of player Bs who announce to select  $r$  actually do so, while any other message makes the likelihood of choosing  $r$  fall by 42 percentage points ( $p=0.045$  after testing  $H_0 : CT\_ReassMess = CT\_NonReassMess$ ) – only 39% (57%) of those sending message "I will choose  $l$ " ("I will choose either  $l$  or  $r$ ") selects  $r$ . Similarly, reputation is not a perfect predictor of reliability. Nonetheless, 94% of player Bs

<sup>35</sup>Note, this way of separating player Bs implies that the first group gathers only those players that constantly played  $r$  before the current interaction. As a result, any player B with a perfect record who chooses  $l$  once in the game drops out from this category permanently, and becomes  $B_{IP}$  ever since.

Table 1.13: Statistical support to Table 1.12

	Reliant A Pr( $R$ )		Reliable B Pr( $r$ )		Cooperation Pr( $R, r$ )		Coordination Pr( $(R, r) \cup (L, l)$ )		Type I errors Pr( $L, r$ )		Type II errors Pr( $R, l$ )	
	coef	p-val.	coef	p-val.	coef	p-val.	coef	p-val.	coef	p-val.	coef	p-val.
Intercept	0.490	0.000	0.807	0.000	0.413	0.000	0.530	0.000	0.393	0.000	0.077	0.000
CT_ReassMess	0.232	0.009	0.096	0.249	0.243	0.019	0.157	0.038	-0.147	0.039	-0.011	0.813
CT_NonReassMess	-0.298	0.004	-0.327	0.002	-0.290	0.006	0.045	0.362	-0.037	0.521	-0.008	0.608
OT_Round1	-0.323	0.001	-0.040	0.682	-0.380	0.001	-0.397	0.000	0.340	0.000	0.057	0.486
OT_PerfRep	0.287	0.049	0.133	0.067	0.343	0.033	0.267	0.036	-0.211	0.062	-0.056	0.008
OT_ImPerfRep	-0.170	0.025	-0.118	0.242	-0.176	0.048	-0.063	0.112	0.057	0.159	0.005	0.840

**Note.** In each column, we present the results of session-clustered (9 clusters in total, 100 observations per cluster) OLS regressions of treatment-related variables on player A's decision  $R$ , player B's decision  $r$ , cooperative outcome  $(R, r)$ , coordinated outcome  $(R, r) \cup (L, l)$ , Type I error  $(L, r)$  and Type II error  $(R, l)$  (columns 1-6, respectively). The intercept represents the frequency in the baseline treatment, the coefficient of each explanatory variable is interpreted as a change induced by a given signal: reassuring (CT\_ReassMess) and non-reassuring (CT\_NonReassMess) message in the Communication Treatment; perfect (OT\_PerfRep) and imperfect (OT\_ImPerfRep) reputation in the Observation Treatment; OT\_Round1 controls for the effect of the first round of the Observation Treatment, where no signals were available yet.

entering an interaction with a full record of weakly dominant actions continue to behave this way, while only 69% of the subjects with an imperfect reputation select  $r$  ( $p < 0.001$  after testing  $H_0 : OT\_PerfRep = OT\_ImPerfRep$ ). Thus, in both treatments positive signals provide an accurate screening of player Bs' intentions. The rate of reliable partners is much higher among those delivering such a signal – through either the message announcing a play  $r$  or a perfect reputation – than others. As compared to the baseline treatment, the reliability rate amongst player Bs announcing decision  $r$  in the communication treatment is slightly improved with respect to the baseline ( $p=0.249$ ), and significantly decreases among player Bs sending one of the two other messages ( $p=0.002$ ). The reliability rate among players with a perfect reputation in the observation treatment is also improved with respect to the baseline, this time significantly ( $p = 0.067$ ), while an imperfect reputation is only slightly detrimental in this respect ( $p=0.242$ ).

Furthermore, player As appear to account for this information by relying more on player Bs delivering a positive signal: from 49% in the baseline, the reliance rate increases up to 72% against a reassuring message and 78% against perfect reputation. As a result, any positive signal induces a significant increase in the rate of cooperation and coordination, along with a fall in type I errors (based on the corresponding regression models in Table 1.13, all comparisons with the baseline are significant at the 5% level, with exception of type I error in the observation treatment, for which  $p=0.062$ ). At the same time, both types of negative signals substantially decrease the likelihood of achieving the most efficient outcome  $(R, r)$  (both  $p < 0.050$ ).

When comparing communication to observation, we find that the performance of positive signals is similar in terms of both cooperation and coordination, although we observe quantitatively higher rates following a perfect reputation rather than a reassuring message – 76% for coordination on the most efficient equilibrium following the former, against 67% following the latter; and 80% for coordination on either of the two Nash equilibria following a perfect reputation, against 69%

following a reassuring message.<sup>36</sup>

Interesting differences arise, by contrast, as regards the effect of negative signals. Player Bs with an imperfect reputation are substantially more reliable than those sending a non-reassuring message – the difference in likelihoods of decision  $r$  is about 21 percentage points and statistically significant ( $p=0.045$ ).<sup>37</sup> Although player As take this into account (the analogous difference between the average likelihoods of reliance amounts to 13 percentage points, with  $p=0.093$ ), we nonetheless find that an imperfect reputation deteriorates coordination as compared to the baseline, while non-reassuring messages improve the odds that player A accurately accommodates partner’s true intentions. Comparing the two treatments, the difference in coordination amounts to 11 percentage points ( $p = 0.056$ ).

### The effects of long-term observation

In Table 1.14, we further investigate how subjects use the environment allowing for a long-term observation. We break down player Bs’ reputation in the observation treatment (restricted to rounds 3-10) according to history (H) – the very last observation – and pre-history (PH) – the proportion of decisions  $r$  among all the remaining observations (preceding the last round). For each decision observed recently – either  $l$  or  $r$  – we consider the marginal effect of four intervals of proportions of decisions  $r$  in the remaining rounds:  $[0; 0.5[$ ,  $[0.5; 0.7[$ ,  $[0.7; 0.9[$ ,  $1$ . We find that both components of reputation, historical and pre-historical, are taken into account by player As, and that consistent signals in which both elements point to the same behavior facilitate coordination to a greater extent than more ambiguous ones in which the two components diverge.

When the previous decision is  $r$ , reliance decreases as decisions  $r$  become less frequent in earlier stages (the coefficients of  $H(r) \times PH \in [0; 0.5[$  and  $H(r) \times PH \in [0.5; 0.7[$  are both negative with  $p = 0.015$  and  $p = 0.063$ , respectively). In the other extreme, decision  $l$  in the previous round has a negative but statistically insignificant effect on player A’s behavior when reputation is perfect in earlier stages (the  $p$ -value for the coefficient of  $H(l) \times PH(1)$  is 0.134). A bad pre-historical reputation (lower than 0.5), in turn, induces player As to rely on the most recent decision they observe ( $p = 0.025$ ).<sup>38</sup>

As regards cooperation and coordination, reputations  $H(r) \times PH [0; 0.5[$  and  $H(r) \times PH [0.5; 0.7[$  lead to a substantial decrease in the likelihoods of both outcomes as compared to a perfect reputation  $H(r) \times PH(1)$  (cooperation:  $p=0.029$ ,  $p=0.052$ ; coordination:  $p=0.091$ ,  $p=0.038$ , respectively). For a perfect pre-history, we observe a substantial positive effect of observing the most

---

<sup>36</sup>None of these differences are significant, though: testing  $H_0 : CT\_ReassMess = OT\_PerfRep$  gives  $p=0.475$  for cooperation,  $p= 0.383$  for coordination,  $p=0.572$  and  $p=0.326$  for type I and type II errors, respectively.

<sup>37</sup>All comparisons made in this paragraph are based on tests of  $H_0 : CT\_NonReassMess = OT\_ImPerfRep$  in the regressions in Table 1.13.

<sup>38</sup>These results are obtained through an additional  $t$ -test for equality of coefficients. For instance, in the latter case we test  $H_0 : H(r) \times PH [0; 0.5[ = H(l) \times PH [0; 0.5[$  in regression in column 3.

Table 1.14: Marginal effect of components of past history in the observation treatment

	N	Pr( $R$ )		Pr( $r$ )		Pr( $R, r$ )		Pr[( $R, r$ ) $\cup$ ( $L, l$ )]	
		coef	p-val.	coef	p-val.	coef	p-val.	coef	p-val.
$H(r) \times PH(1)$	125	0.808	0.009	0.960	0.000	0.784	0.012	0.800	0.012
$H(r) \times PH[0.7; 0.9[$	29	-0.291	0.152	-0.132	0.336	-0.336	0.205	-0.248	0.251
$H(r) \times PH[0.5; 0.7[$	28	-0.487	0.063	-0.174	0.139	-0.498	0.052	-0.336	0.038
$H(r) \times PH[0; 0.5[$	17	-0.455	0.015	-0.489	0.068	-0.666	0.029	-0.388	0.091
$H(l) \times PH(1)$	9	-0.586	0.134	-0.182	0.106	-0.673	0.043	-0.578	0.069
$H(l) \times PH[0.7; 0.9[$	8	-0.433	0.060	-0.335	0.208	-0.534	0.148	-0.300	0.249
$H(l) \times PH[0.5; 0.7[$	8	-0.558	0.158	-0.210	0.408	-0.534	0.170	-0.300	0.079
$H(l) \times PH[0; 0.5[$	16	-0.683	0.016	-0.335	0.015	-0.722	0.018	-0.425	0.164

**Note.** The Table summarizes the results – estimated coefficients and corresponding  $p$ -values – of a session-clustered (3 clusters from the Observation Treatment, 80 observations per cluster, standard errors corrected with a delete-one jackknife) OLS regressions of dummies representing player A's decision  $R$ , player B's decision  $r$ , cooperative outcome  $(R, r)$  and coordinative outcome  $(R, r) \cup (L, l)$  (columns 3-6, respectively) on a set of reputation-related variables. Only observations from rounds 3-10 from the observation treatment are included. For instance, variable  $H(r) \times PH[0.5; 0.7[$  stands for a reputation in which the decision from the most recent round (history) is  $r$ , and the proportion of decisions  $r$  in the remaining rounds (pre-history) is greater or equal to 0.5 and less than 0.7.

recent decision  $r$  instead of  $l$ , on the probability of achieving both  $(R, r)$  ( $p=0.043$ ) and  $(R, r) \cup (L, l)$  ( $p=0.069$ ). Furthermore, reputations  $H(l) \times PH[0; 0.5[$  and  $H(l) \times PH[0.5; 0.7[$ , where both pre-history and history are imperfect, happen to provide a better screening of player Bs' true intentions, which increases the likelihood of achieving a coordinated outcome, than a reputation  $H(l) \times PH(1)$ , where only the most recent decision is imperfect ( $p=0.059$ ,  $p=0.079$ , respectively).<sup>39</sup>

### The determinants of reliance

We test the marginal significance of the effects highlighted above by estimating treatment-specific models of the probability of observing a reliant decision from player As. We include individual random effects to account for individual heterogeneity and condition observed behavior on time-dummies, observed past behavior and both group and individual information about player Bs. This conditioning controls for player As' common experience due to interactions with the same player Bs over time. Due to the introduction of lagged variables, all models are estimated on data from rounds 2-10. The results from probit regressions are presented in Table 1.15.

In the baseline treatment, the round dummies indicate small variations over time. This effect of time reflects repetition-based learning about player Bs behavior. The variable *Population\_B* is constructed as the proportion of decisions  $r$  among all decisions made in the entire population of player Bs in the earlier rounds of the experimental session – this measures global variations from on round to the other. The effect of this variable thus measures how the true behavior of

<sup>39</sup>These results are obtained through additional tests for equality of coefficients in the regression in column 6:  $H_0 : H(l) \times PH(1) = H(l) \times PH[0.5; 0.7[$  and  $H_0 : H(l) \times PH(1) = H(l) \times PH[0.7; 0.9[$ , respectively.

Table 1.15: Probit regressions on player As' reliance in the three experimental games

	Baseline			Communication			Observation		
	coef	p-val.	ME	coef	p-val.	ME	coef	p-val.	ME
Intercept	-3.56	0.03	—	-0.97	0.59	—	-1.77	0.12	—
Male_A	0.71	0.16	0.28	0.43	0.16	0.16	0.53	0.11	0.20*
$\mathbf{1}_{Round=3}$	0.35	0.42	0.14	0.00	0.99	0.00	0.37	0.39	0.14
$\mathbf{1}_{Round=4}$	-0.06	0.90	-0.02	-0.04	0.93	-0.01	0.28	0.58	0.10
$\mathbf{1}_{Round=5}$	0.07	0.88	0.03	0.17	0.68	0.06	0.00	0.99	0.00
$\mathbf{1}_{Round=6}$	-0.24	0.66	-0.10	0.07	0.88	0.03	-0.42	0.41	-0.16
$\mathbf{1}_{Round=7}$	-0.75	0.21	-0.27	-0.68	0.18	-0.26	-0.01	0.98	-0.01
$\mathbf{1}_{Round=8}$	-1.12	0.08	-0.38**	-0.24	0.65	-0.09	-0.17	0.79	-0.07
$\mathbf{1}_{Round=9}$	-1.12	0.11	-0.38**	-1.03	0.08	-0.39*	-0.69	0.30	-0.27
$\mathbf{1}_{Round=10}$	-2.07	0.01	-0.54***	-0.97	0.14	-0.37	-1.08	0.13	-0.41*
$\mathbf{1}_{Trusted\ Before}$	0.33	0.60	0.13	0.15	0.74	0.06	-0.10	0.83	-0.04
$\times Round \times Ratio\_Rr$	0.28	0.02	0.11**	0.23	0.01	0.08***	0.24	0.01	0.09***
Population_B	3.20	0.08	1.28*	-1.05	0.55	-1.28	-0.47	0.93	-0.05
Message_r	—	—	—	1.75	0.00	0.62***	—	—	—
Message_l	—	—	—	-0.30	0.46	-0.12	—	—	—
Reputation $\in (0.5; 0.6]$	—	—	—	—	—	—	-0.71	0.42	-0.28
Reputation $\in (0.6; 0.7]$	—	—	—	—	—	—	-0.03	0.95	-0.01
Reputation $\in (0.7; 0.8]$	—	—	—	—	—	—	0.19	0.66	0.07
Reputation $\in (0.8; 0.9]$	—	—	—	—	—	—	0.92	0.03	0.29**
Perfect_Reputation	—	—	—	—	—	—	1.92	0.00	0.66***
N	270			270			270		
Log L	-110.68			-117.73			-120.77		

**Legend.** \*, \*\*, \*\*\* indicate statistical significance at the 10%, 5%, 1% level.

**Note.** Panel probit regressions with individual random effects. The dependent variable is a dummy indicating whether player A makes a reliant decision  $R$ , observations come from rounds 2-10. Marginal effects (ME) are calculated under the assumption that the individual factors  $u_i$  are equal to 0. Male\_A indicates whether player A is a male. The variables  $\mathbf{1}_{Round=3}, \dots, \mathbf{1}_{Round=10}$  are round fixed effects – round 2 is omitted.  $\mathbf{1}_{Trusted\ Before}$  is a dummy variable switched to one once player A has chosen  $R$  in the past. *Population\_B* is computed in each round (accept for round 1) for every experimental session as the proportion of decisions  $r$  among all past decisions made by the population of player Bs. For the communication treatment (middle part), the flow of information is included through the content of received message ("I will choose  $r$ " or "I will choose  $l$ "). For the observation treatment (right-hand side) the reputation of the current interaction partner is accounted for through dummies for each 10% range (past reliability less than or equal to 50% is the reference).

the population of player Bs is accommodated for in the experiment – although this information is never available as such to player As in this treatment. Both the coefficient and the marginal effect are significant. In the baseline treatment, repetition-based learning induces a positive correlation between reliance and the overall rate of reliable partners in the population of player Bs. Interestingly, this effect vanishes in the two information-enhancing treatments. Once any information flow becomes available (through either cheap-talk or observation) trustworthiness becomes driven only by the specific signal.

In the communication treatment, the information appears to rely solely on the reassuring message ("I will choose  $r$ ") which substantially improves the odds of taking action  $R$ . Message "I will choose  $l$ " does not significantly change the probability of decision  $R$ , as compared to an uninformative message. For the observation treatment, we discretize the reputation observed by

Table 1.16: Average payoffs from the use of dominant strategies

	Reliability rate			Average payoffs					
	$Pr(B_P)$	$Pr(r B)$		$E(Payoff(B) r)$			$E(Payoff(A) R)$		
		$B_P$	$B_{IP}$	$B_P$	$B_{IP}$	$B$	$B_P$	$B_{IP}$	$B$
Baseline	55.6%	92.0%	66.7%	4.07€	4.08€	4.07€	9.56€	7.93€	8.85€
Communication	53.0%	92.3%	66.1%	4.42€	4.21€	4.34€	9.78€	8.41€	9.23€
— Following m(r)	62.1%	97.6%	77.9%	4.48€	4.50€	4.49€	9.85€	8.63€	9.39€
Observation	54.8%	93.9%	68.9%	4.61€	3.69€	4.26€	9.82€	8.21€	9.41€

**Note.** In each column, outcomes are conditioned on interacting with player Bs ( $B$ ), who are then separated into two subgroups, depending on whether they enter the interaction with a perfect reputation ( $B_P$ ) or not ( $B_{IP}$ ). Due to such conditioning, data come from rounds 2-10. For each treatment in rows, the left-hand side reports the reliability rate in the main population player Bs as well as in each of the subgroups. In the right hand-side, the middle column provides the average payoff earned by player Bs who choose the payoff maximizing decision  $r$ . The last column provides the average payoff earned by player As who rely on their partner by choosing  $R$ .

player A into 6 intervals, and estimate separately the effect of perfect reputation. As compared to the reference category of a bad reputation (past reliability not exceeding 50%), only very high levels of reputation (over 80% of decisions  $r$  in the past) are able to induce a significant increase in reliance. Partners' perfect reputation induces the highest increase in the use of a reliant strategy, corresponding to a 66 percentage points increase in the likelihood of a reliance.

In all games, player As who receive positive signals (through the repetition of game in the baseline treatment, reassuring messages in the communication treatment and the quality of reputation in the observation treatment) are significantly more willing to rely on partners. One virtue of this change in behavior is to strengthen efficiency through reinforcement-based learning. The dummy variable  $\mathbf{1}_{Trusted\ Before}$  indicates whether player A has already relied on her partner in the past. This variable *per se* is a very poor predictor of the likelihood of current reliance. We capture reinforcement through interacting this variable with time (variable *Round*) and a measure of success, *Ratio\_Rr*, constructed as the proportion of outcomes  $(R, r)$  among all decisions  $R$ . In all regressions, estimated coefficients indicate that decision  $R$  becomes more likely over time the more fruitful are the historical attempts to rely on partners. This reinforcement-based learning is significant and has comparable marginal effects in all treatments.

### Empirical efficiency of reliance

Table 1.16 summarizes the average payoffs earned by and against player Bs with different reputation when the decisions of the payoff maximizing strategy,  $r$  for player B and  $R$  for player A, are actually played.<sup>40</sup> As compared to the baseline treatment, the introduction of either communication or observation does not change a lot the average welfare of trustful player As, while

<sup>40</sup>Due to this conditioning on historical behavior, we use observations from rounds 2-10.



substantially benefiting player Bs. The welfare of reliant player As increases, though, when reliance is conditioned on either of the positive signals – conveyed through a reassuring message or an observed perfect reputation. Conditional on receiving a positive signal, observation happens to be more welfare improving for both players than communication. Keeping with a message announcing decision  $r$  brings on average 4.49€ to reliable player Bs, while observed perfect reputation yields on average 4.61€; for reliant player As, the average payoff increases from 9.39€ to 9.82€, respectively.

From player As' perspective, reliance is empirically dominated in all three experimental conditions. Recall that player A can always guarantee 9.75€ in the game by choosing  $L$ , in which case player B experiences a 3€ payoff. As compared to this threshold, reliance never dominates the secure choice: the average earnings are 8.85€ in the baseline, 9.23€ in the communication treatment and 9.41€ in the observation treatments. These average payoffs closely reflect a common lack of reliability in the population of player Bs. In all treatments player Bs' past behavior is strongly related to their current intentions, in such a way that the emergence of two distinct types of player Bs – already reported in the context of observation treatment – extends to other experimental conditions. This is accounted for in the table by distinguishing highly reliable  $B_P$  players (who always played  $r$  in the past and whose likelihood of choosing  $r$  is 92%-94%) and less reliable  $B_{IP}$  players (who deviated from playing  $r$  in the past, and whose probability of choosing  $r$  is 66%-69%).<sup>41</sup>

Although in the communication treatment player As cannot observe player Bs' history and directly distinguish between the two types, a reassuring message conveys an indirect (and imperfect) information about partner's type. A strong majority (62%) of subjects sending a reassuring message are highly reliable  $B_P$  players who then choose  $r$  97.6% of times and provide an average gain of 9.85€ to a reliant partner (as compared to 77.9% and 8.63€, respectively, observed among  $B_{IP}$  players who send this signal). Consequently, the presence of weakly reliable player Bs severely undermine the average gain in welfare due to relying on a partner sending a reassuring message which ultimately makes unreliance empirically more beneficial.

Observation, by contrast, allows player As to identify both types of player Bs, to reward a good reputation, and to punish for a bad reputation. This results in a strong inter-type gap in player Bs' average gains (4.61€ for  $B_P$  subjects vs. 3.69€ for  $B_{IP}$  subjects). Although a large difference in player As' average payoffs from relying on both types of player Bs (9.82€ for  $B_P$  subjects vs. 8.21€ for  $B_{IP}$  subjects) justifies such a discriminatory behavior, player As do not appear to be conservative enough, since the average gain from action  $R$  (9.41€) remains lower than the value of safe option.

---

<sup>41</sup>The differences in reliability between both types of player Bs, shown in the second column of Table 1.16, are statistically significant according to our parametric test for equality of proportions, with  $p$ -values equal to 0.052, 0.002, 0.002, 0.050 for subsequent rows.

## Discussion

The principle findings established in this section are twofold. First, we empirically demonstrate that coordination failures in a dominance-solvable  $2 \times 2$  coordination game based on Rosenthal (1981) have a different structure than what is suggested by previous experimental studies. In all previous experiments, player Bs' actions are elicited conditional on player As being reliant. Given that player As seldom rely on their partners, and that the proportion of own-payoff-maximizers among those relied upon is generally high, these studies conclude that player As often fail to rely on player Bs' own-payoff-maximization. This indeed happens in the cases where player A interacts with a perfect own-payoff-maximizer and nonetheless chooses  $L$  (what we call a *type I error*, following the idea that player A's decision provides a test of player B's preferences). Yet, conditioning the observability of player Bs' choices on player As' prior reliance makes it impossible to judge whether action  $L$  is virtually justified or not. The only coordination failure this design might capture is *type II error*, in which player A relies on a player B who then uses the weakly dominated strategy. In order to fully account for the occurrence of coordination failures in the game, one thus needs to observe not only player As' degree of unreliance, but also player Bs' behavior unconditional on what player As do. Our experiment implements the normal form of the game, so that decisions are elicited from both players in each interaction. Thanks to this design, we show that the empirical puzzle raised by Rosenthal's conjecture is more sophisticated than previously thought. Although type I errors are more widespread than type II errors, unreliance is often justified. This occurs because unreliability (through decision  $l$ ) is widespread among player Bs, and insensitive to the treatment variables. In this sense, our study provides genuine evidence that a relatively weak degree of reliance from player As may be rational given the behavior of the population of player Bs. Coordination failures due to the strategic uncertainty faced by player As still prevail, however, since a lot of opportunities to reach the efficient outcome are missed. This is precisely what enhanced information seeks to resolve.

Second, we observe that repetition-based learning stimulates reliance which, in turn, increases the frequency of the most efficient outcome. However, this improvement comes at a price in terms of coordination failures: since player As learn about the general population of player Bs rather than their actual interaction partner, more attempts to rely on the payoff-maximizing behavior fail. Any kind of subject-specific information, through cheap-talk or observation, appears as a substitute to repetition-based learning. Both kinds of positive signals (through either an announcement to play the weakly dominant strategy or a reputation to have done so frequently in the past) provide an accurate screening of player Bs' intentions and induce a rise in player As' reliance. Communication appears more informative than observation, in turn, if signals are negative: a non-reassuring message allows for a better strategic coordination than an imperfect reputation. This confirms previous findings that information-transmission mechanisms affect both the frequency and the form of coordination failures (Parkhurst, Shogren, and Bastian (2004)).

Many experimental studies of communication and observation in coordination games report that such institutions help overcome coordination failures and foster efficient cooperation (see, *e.g.*, Cooper, DeJong, Forsythe, and Ross (1992), Crawford (1998), Charness (2000), Duffy and Feltovich (2002), Duffy and Feltovich (2006), Blume and Ortmann (2007)). To our best knowledge, we are the first to study this kind of mechanisms in an *asymmetric* coordination problem, in which agents' interests are aligned, but the stakes and the strategic risks they face differ. An advantage of our design as compared to symmetric games is that players' functions in the information-transmission process (that is, being either the sender or the receiver) are a natural consequence of their strategic position in the game, rather than an outcome of a random draw.

In sharp contrast to previous experimental results, we report that although communication and observation restrain coordination failures, they do not importantly improve the overall efficiency of outcomes, which happens mostly due to the persistence of player Bs' inefficient behavior. An important question raised by this experiment is the following: which features of decision-making environment give rise to this behavior? This issue is addressed in the following section.

## 1.4 Summary and conclusion

Both experiments presented in this Chapter revisit the Beard and Beil (1994) two-player coordination game with two Nash equilibria: one is Pareto-efficient, the other is Pareto-inefficient and involves a weakly dominated strategy. Existing experiments using this game robustly show that suboptimal outcomes arise as a result of two puzzling behaviors: *(i)* subjects doubt that the other players will seek to maximize their own payoff and *(ii)* these doubts are, in some instances, justified.

Here, we first report on new experiments investigating whether the inequality in payoffs between players, maintained in most lab implementations of this game, may explain such behavior. Our data clearly show that the failure to maximize personal payoffs, as well as the fear that others might act this way, do not stem from inequality aversion. This result is robust to: varying saliency of decisions, repetition-based learning and cultural differences between France and Poland.

Second, we assess whether information about the interaction partner helps eliminate inefficiency in this game. Our treatments involve three information-enhancing mechanisms: repetition and two kinds of individual signals, messages from partner or observation of his past choices. Repetition-based learning increases the frequencies of the most efficient outcome and the most costly strategic mismatch. Moreover, it is superseded by individual signals. Like previous empirical studies, we report that signals provide a screening of partners' intentions that reduces the frequency of coordination failures. Unlike these studies, we find that the transmission of information between partners, either via messages or observation, does not suffice to significantly increase the overall efficiency of outcomes. This happens mostly because information does not restrain the use of the dominated action by senders.

## 1.5 Supplementary material

### Supplementary material for Section 1.2

#### Written instructions

*Authors' note: Experimental instructions for the Baseline Treatment 1. In other treatments, we use identical instructions with payoffs modified accordingly.*

You are about to take part in an experiment in which you can earn money. The amount of your gains will depend on your decisions, as well as on decisions made by other participants.

Before starting, we would like to ask you to answer a few standard questions (concerning your age, education, profession, ...) which will help us to get to know you better. **This information, as well as the amount of your gains from this experiment, will remain strictly confidential and anonymous.**

Please, fill in the questionnaire using the interface on your computer screen, which is divided into three parts:

- In the *top* section, you will find information that might help you in making decisions.
- In the *middle* section, you will submit your decisions by clicking on a relevant button.
- In the *bottom* section, you will see all your decisions and gains from previous rounds of the experiment.

**Thank you.**

## THE EXPERIMENT

The experiments consists of several identical rounds. In each round, participants are divided by groups of two. Each pair consists of one player A and one player B. You will be randomly assigned to your role — player A or player B — at the beginning of the experiment, and retain it throughout the experimental session. A message on your computer screen will inform you about your role. **Your role will remain unchanged throughout the entire experiment.**

### WHAT HAPPENS IN EACH ROUND

At the beginning of each round, participants are be matched into pairs: if your are player A, then a player B is randomly selected to your complete pair; analogously, if your are player B, then a player A is randomly selected to complete your pair. Your pair will **change after each round**, and two participants in opposite roles **may interact at most once during the experiment.**

Each round consists of 4 stages.

**Stage 1.** At the beginning, a participant is randomly matched to your group.

**Stage 2.** Player A chooses between  $L$  and  $R$  by clicking on a relevant button on his/her computer screen.

**Stage 3.** Player B chooses between  $l$  and  $r$  by clicking on a relevant button on his/her computer screen.

**Stage 4.** End of the round and each player is informed about his/her earnings:

- If **player A chose  $L$** , then **regardless of player B's decision**:
  - ▶ Player A earns 9.75 € in this round;
  - ▶ Player B earns 3 € in this round;
- If **player A chose  $R$**  then:
  - if **player B chose  $l$** :
    - ▶ Player A earns 3 € in this round;
    - ▶ Player B earns 4.75 € in this round;
  - if **player B chose  $r$** :
    - ▶ Player A earns 10 € in this round;
    - ▶ Player B 5 € in this round;

At the end of each round, a message on your computer screen will inform you that either a new round is about to start, or that the experiment ends.

## **PAYMENT OF YOUR EARNINGS**

At the end of the experiment, **one round is picked at random**. Each participant receives a sum in EUR corresponding to the amount he/she earned in this round, plus a bonus of 5 € for completing the experiment. Payments are made individually and in cash.

For obvious reasons, **you are not allowed to talk during the experiment**. Participants who violate this rule will be excluded from the experiment and all payments. It is crucial that you understand perfectly the rules of this experiment. Should you have any questions to ask, please raise your hand.

**Thank you for your participation.**

### Replication of the model in Table 1.16 with treatment-specific dummies

This section reports the results from a regression similar to the one presented in Section 1.2.3 (Table 1.7), in which the explanatory variables include treatment-specific dummies instead of design-specific dummies. The results confirm the robustness of the treatment effects derived from the 2-by-2 mean comparisons.

	Pr(R)				Pr(r)			
	coef	p-val.	ME	p-val.	coef	p-val.	ME	p-val.
<i>Intercept</i>	-1.241	0.040	—	—	-0.284	0.572	—	—
<i>BT1</i>	0.091	0.658	0.018	0.649	0.107	0.774	0.019	0.772
<i>ET3</i>	0.055	0.803	0.011	0.799	0.105	0.770	0.018	0.768
<i>ET4</i>	0.396	0.076	0.077	0.052	0.201	0.544	0.035	0.545
<i>ET2</i>	0.566	0.084	0.110	0.059	0.662	0.096	0.114	0.093
<i>BT2</i>	0.413	0.089	0.080	0.063	0.635	0.062	0.110	0.064
<i>Poland</i>	0.057	0.774	0.011	0.771	0.171	0.431	0.030	0.427
Learning: Player A								
$\mathbf{1}[R_{t-1}]$	2.172	0.000	0.422	0.000	—	—	—	—
Population $B_{t-1}$	0.332	0.657	0.065	0.655	—	—	—	—
Learning: Player B								
$\mathbf{1}[r_{t-1}]$	—	—	—	—	1.412	0.000	0.244	0.000
Population $A_{t-1}$	—	—	—	—	-0.153	0.853	-0.027	0.851
<b>Round Dummies</b>								
$t = 3$	-0.022	0.926	-0.004	0.925	0.044	0.729	0.008	0.728
$t = 4$	-0.126	0.505	-0.024	0.499	0.064	0.650	0.012	0.652
$t = 5$	-0.148	0.296	-0.028	0.281	0.011	0.950	0.002	0.950
$t = 6$	-0.141	0.435	-0.027	0.435	0.047	0.753	0.009	0.752
$t = 7$	-0.073	0.559	-0.014	0.550	0.154	0.543	0.027	0.535
$t = 8$	-0.270	0.075	-0.052	0.052	0.364	0.080	0.058	0.074
$t = 9$	-0.092	0.539	-0.018	0.535	-0.011	0.948	-0.002	0.947
$t = 10$	-0.192	0.266	-0.037	0.262	0.165	0.398	0.029	0.383

## Overview of the results from all treatments

This section provides a full summary of round-by-round outcomes from all treatments. These data suggest that aggregate outcomes are fairly stable over time; all changes only occur at the very beginning of the experiment, mostly between the first and the second round.

Game	Decisions		Coordination		Failure		Decisions		Coordination		Failure	
	$R$	$r$	$(R, r)$	$(L, l)$	$(L, r)$	$(R, l)$	$R$	$r$	$(R, r)$	$(L, l)$	$(L, r)$	$(R, l)$
	<b>Round 1</b>						<b>Round 2</b>					
BT1	0.233	0.800	0.233	0.200	0.567	0.000	0.367	0.800	0.300	0.133	0.500	0.067
ET1	0.467	0.633	0.300	0.200	0.333	0.167	0.467	0.700	0.333	0.167	0.367	0.133
ET3	0.575	0.800	0.525	0.150	0.275	0.050	0.550	0.875	0.500	0.075	0.375	0.050
ET4	0.633	0.767	0.467	0.067	0.300	0.167	0.733	0.800	0.600	0.067	0.200	0.133
BT2	0.500	0.933	0.500	0.067	0.433	0.000	0.633	0.900	0.567	0.033	0.333	0.067
ET2	0.540	0.840	0.420	0.040	0.420	0.120	0.720	0.860	0.600	0.020	0.260	0.120
	<b>Round 3</b>						<b>Round 4</b>					
BT1	0.500	0.900	0.467	0.067	0.433	0.033	0.500	0.833	0.467	0.133	0.367	0.033
ET1	0.433	0.767	0.300	0.100	0.467	0.133	0.533	0.733	0.400	0.133	0.333	0.133
ET3	0.600	0.825	0.500	0.075	0.325	0.100	0.600	0.800	0.450	0.050	0.350	0.150
ET4	0.767	0.733	0.567	0.067	0.167	0.200	0.767	0.800	0.633	0.067	0.167	0.133
BT2	0.767	0.900	0.700	0.033	0.200	0.067	0.733	0.967	0.700	0.000	0.267	0.033
ET2	0.740	0.920	0.660	0.000	0.260	0.080	0.760	0.940	0.720	0.020	0.220	0.040
	<b>Round 5</b>						<b>Round 6</b>					
BT1	0.600	0.733	0.500	0.167	0.233	0.100	0.600	0.800	0.500	0.100	0.300	0.100
ET1	0.467	0.733	0.300	0.100	0.433	0.167	0.400	0.667	0.300	0.233	0.367	0.100
ET3	0.575	0.800	0.475	0.100	0.325	0.100	0.650	0.850	0.525	0.025	0.325	0.125
ET4	0.700	0.767	0.567	0.100	0.200	0.133	0.733	0.867	0.667	0.067	0.200	0.067
BT2	0.767	0.967	0.733	0.000	0.233	0.033	0.800	0.867	0.667	0.000	0.200	0.133
ET2	0.800	0.980	0.780	0.000	0.200	0.020	0.780	0.960	0.760	0.020	0.200	0.020
	<b>Round 7</b>						<b>Round 8</b>					
BT1	0.567	0.767	0.467	0.133	0.300	0.100	0.533	0.767	0.367	0.067	0.400	0.167
ET1	0.500	0.667	0.333	0.167	0.333	0.167	0.433	0.767	0.367	0.167	0.400	0.067
ET3	0.575	0.900	0.525	0.050	0.375	0.050	0.550	0.850	0.475	0.075	0.375	0.075
ET4	0.733	0.833	0.567	0.000	0.267	0.167	0.733	0.933	0.667	0.000	0.267	0.067
BT2	0.800	0.967	0.767	0.000	0.200	0.033	0.800	1.000	0.800	0.000	0.200	0.000
ET2	0.860	0.980	0.840	0.000	0.140	0.020	0.820	1.000	0.820	0.000	0.180	0.000
	<b>Round 9</b>						<b>Round 10</b>					
BT1	0.567	0.800	0.500	0.133	0.300	0.067	0.433	0.867	0.333	0.033	0.533	0.100
ET1	0.367	0.733	0.267	0.167	0.467	0.100	0.500	0.867	0.400	0.033	0.467	0.100
ET3	0.550	0.850	0.475	0.075	0.375	0.075	0.525	0.725	0.375	0.125	0.350	0.150
ET4	0.767	0.833	0.633	0.033	0.200	0.133	0.733	0.900	0.667	0.033	0.233	0.067
BT2	0.833	0.900	0.767	0.033	0.133	0.067	0.800	1.000	0.800	0.000	0.200	0.000
ET2	0.860	0.960	0.820	0.000	0.140	0.040	0.880	0.920	0.800	0.000	0.120	0.080

## Supplementary material for Section 1.3

### Written instructions

*Authors' note: Experimental instructions for the Baseline Treatment. Additional parts that were only included in the instructions for the Communication Treatment (CT) or the Observation Treatment (OT) appear in italics preceded by a comment "In CT only:" or "In OT only:". The stage numbers vary from one treatment to the other.*

You are about to take part in an experiment in which you can earn money. The amount of your gains will depend on your decisions, as well as on decisions made by other participants.

Before starting, we would like to ask you to answer a few standard questions (concerning your age, education, profession, ...) which will help us to get to know you better. **This information, as well as the amount of your gains from this experiment, will remain strictly confidential and anonymous.**

Please, fill in the questionnaire using the interface on your computer screen, which is divided into three parts:

- In the *top* section, you will find information that might help you in making decisions.
- In the *middle* section, you will submit your decisions by clicking on a relevant button.
- In the *bottom* section, you will see all your decisions and gains from previous rounds of the experiment.

**Thank you.**

## THE EXPERIMENT

The experiments consists of several identical rounds. In each round, participants are divided by groups of two. Each pair consists of one player A and one player B. You will be randomly assigned to your role — player A or player B — at the beginning of the experiment, and retain it throughout the experimental session. A message on your computer screen will inform you about your role. **Your role will remain unchanged throughout the entire experiment.**

### WHAT HAPPENS IN EACH ROUND

At the beginning of each round, participants are be matched into pairs: if your are player A, then a player B is randomly selected to your complete pair; analogously, if your are player B, then a player A is randomly selected to complete your pair. Your pair will **change after each round**, and two participants in opposite roles **may interact at most once during the experiment**.

Each round consists of 4 stages. *CT: 6 stages; OT: 5 stages*



**Stage 1.** At the beginning, a participant is randomly matched to your group.

**In CT only:**

**Stage 2.** *Player B is asked to send a message to player A by choosing one of the options displayed on his/her computer screen and submitting it by clicking 'OK'. This message does not affect neither players' earnings.*

**Stage 3.** *Player A reads the message from player B and then clicks 'OK' in order to proceed to the next stage.*

**In OT only:**

**Stage 2.** *Player A is informed about the decisions player B has made in all the previous rounds of the experiment. Of course, in the first round no information is available to player A.*

**Stage 4.** Player A chooses between  $L$  and  $R$  by clicking on a relevant button on his/her computer screen.

**Stage 5.** Player B chooses between  $l$  and  $r$  by clicking on a relevant button on his/her computer screen.

**Stage 6.** End of the round and each player is informed about his/her earnings:

- If **player A** chose  $L$ , then **regardless of player B's decision**:
  - ▶ Player A earns 9.75 € in this round;
  - ▶ Player B earns 3 € in this round;
- If **player A** chose  $R$  then:
  - if **player B** chose  $l$ :
    - ▶ Player A earns 3 € in this round;
    - ▶ Player B earns 4.75 € in this round;
  - if **player B** chose  $r$ :
    - ▶ Player A earns 10 € in this round;
    - ▶ Player B 5 € in this round;

At the end of each round, a message on your computer screen will inform you that either a new round is about to start, or that the experiment ends.

## **PAYMENT OF YOUR EARNINGS**

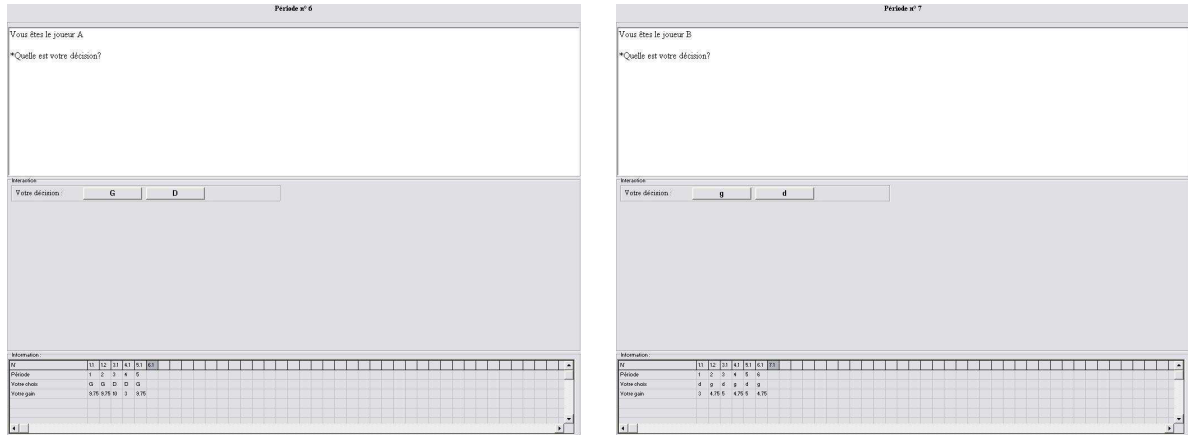
At the end of the experiment, **one round is picked at random**. Each participant receives a sum in EUR corresponding to the amount he/she earned in this round, plus a bonus of 5 € for completing the experiment. Payments are made individually and in cash.

For obvious reasons, **you are not allowed to talk during the experiment**. Participants who violate this rule will be excluded from the experiment and all payments. It is crucial that you understand perfectly the rules of this experiment. Should you have any questions to ask, please raise your hand.

**Thank you for your participation.**

## Screen captures of the experiment

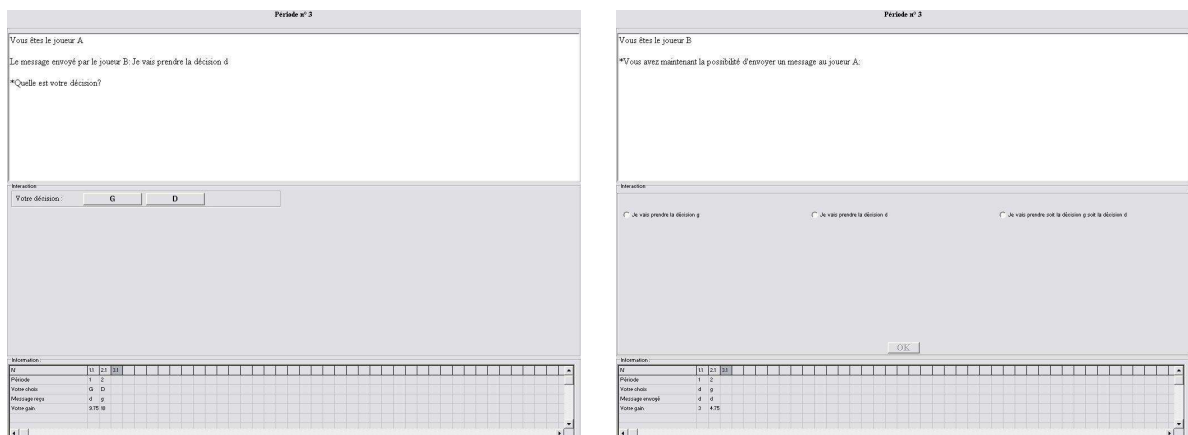
Figure 1.1: Baseline game



(a) Player A

(b) Player B

Figure 1.2: Communication



(a) Player A

(b) Player B

Figure 1.3: Observation

Periode n° 6

Vous êtes le joueur A

Vous allez voir maintenant les décisions du joueur B des périodes précédentes:

Dans la période n° 1 le joueur B a choisi d

Dans la période n° 2 le joueur B a choisi g

Dans la période n° 3 le joueur B a choisi g

Dans la période n° 4 le joueur B a choisi g

Dans la période n° 5 le joueur B a choisi d

Situation

OK

Information

P	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Période	1	2	3	4	5															
Choix du joueur B	g	d	g	g	d															
Score du joueur A	0.75	0	0.75	0.75	0															

(a) Player A

Periode n° 3

Vous êtes le joueur B

\*Quelle est votre décision?

Information

Votre décision :

g

d

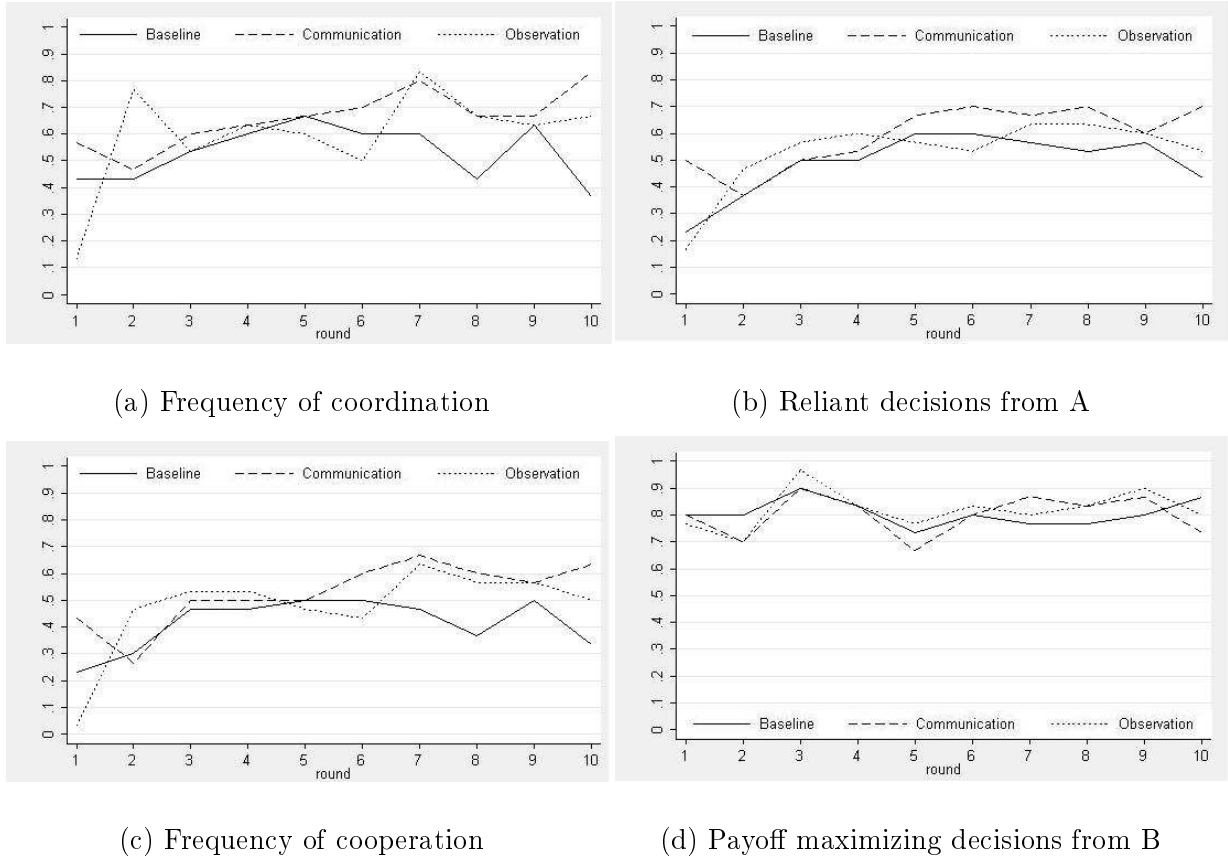
Information

P	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Période	1	2																		
Choix du joueur B																				
Score du joueur B																				

(b) Player B

## Detailed data

Figure 1.4: Evolution of outcomes over rounds



**Note.** Each figure provides the evolution over rounds (*in abscissa*) of the frequency of Nash equilibria (*Fig.a*), the frequency of the Pareto-Nash equilibrium (*Fig.c*), the proportion of reliant decisions from players A (*Fig.b*) and the proportion of  $r$  decisions by players B (*Fig.d*) in the baseline game (Baseline), the pre-play communication treatment (Communication) and the treatment with historical information (Observation).

Table 1.17: Baseline game: evolution of frequencies of choices and outcomes over time

round	Decision $R$	Decision $r$	$(R, r)$	$(R, r) \cup (L, l)$	$(L, r)$	$(R, l)$
1	0.23	0.80	0.23	0.43	0.57	0.00
2	0.37	0.80	0.30	0.43	0.50	0.07
3	0.50	0.90	0.47	0.53	0.43	0.03
4	0.50	0.83	0.47	0.60	0.37	0.03
5	0.60	0.73	0.50	0.67	0.23	0.10
6	0.60	0.80	0.50	0.60	0.30	0.10
7	0.57	0.77	0.47	0.60	0.30	0.10
8	0.53	0.77	0.37	0.43	0.40	0.17
9	0.57	0.80	0.50	0.63	0.30	0.07
10	0.43	0.87	0.33	0.37	0.53	0.10
Total	0.49	0.81	0.41	0.53	0.39	0.08

Table 1.18: Communication game: evolution of frequencies of choices and outcomes over time

round	Decision $R$	Decision $r$	$(R, r)$	$(R, r) \cup (L, l)$	$(L, r)$	$(R, l)$
1	0.50	0.80	0.43	0.57	0.37	0.07
2	0.37	0.70	0.27	0.47	0.43	0.10
3	0.50	0.90	0.50	0.60	0.40	0.00
4	0.53	0.83	0.50	0.63	0.33	0.03
5	0.67	0.67	0.50	0.67	0.17	0.17
6	0.70	0.80	0.60	0.70	0.20	0.10
7	0.67	0.87	0.67	0.80	0.20	0.00
8	0.70	0.83	0.60	0.67	0.23	0.10
9	0.60	0.87	0.57	0.67	0.30	0.03
10	0.70	0.73	0.63	0.83	0.10	0.07
Total	0.59	0.80	0.53	0.66	0.27	0.07

Table 1.19: Observation game: evolution of frequencies of choices and outcomes over time

round	Decision $R$	Decision $r$	$(R, r)$	$(R, r) \cup (L, l)$	$(L, r)$	$(R, l)$
1	0.17	0.77	0.03	0.13	0.73	0.13
2	0.47	0.70	0.47	0.77	0.23	0.00
3	0.57	0.97	0.53	0.53	0.43	0.03
4	0.60	0.83	0.53	0.63	0.30	0.07
5	0.57	0.77	0.47	0.60	0.30	0.10
6	0.53	0.83	0.43	0.50	0.40	0.10
7	0.63	0.80	0.63	0.83	0.17	0.00
8	0.63	0.83	0.57	0.67	0.27	0.07
9	0.60	0.90	0.57	0.63	0.33	0.03
10	0.53	0.80	0.50	0.67	0.30	0.03
Total	0.53	0.82	0.47	0.60	0.35	0.06

## Chapter 2

# Coordination with communication under oath

*Based on joint work with Nobuyuki Hanaki, Nicolas Jacquemet, Stéphane Luchini and Jason Shogren.*<sup>1</sup>

### 2.1 Introduction

A coordination game captures the idea that value can be created when people coordinate their noncooperative actions in a strategic environment (see Schelling, 1960; Cooper, DeJong, Forsythe, and Ross, 1990). If people coordinate their otherwise sovereign actions, they can achieve the first best equilibrium among many suboptimal ones. Coordination failure arises when people fail to realize the first best outcome due to strategic uncertainty – the risk associated with not knowing how your opponent will play the game (see, *e.g.*, the survey by Devetag and Ortmann, 2007). Better communication between players is the most frequently prescribed institution to overcome coordination failure. Accumulated evidence shows *(i)* communication can improve efficiency; but *(ii)* many coordination failures still remain because the messages are non-binding cheap talk, *i.e.*, there are no real economic consequences to players who do not act in accordance to the message (see, *e.g.*, Cooper, DeJong, Forsythe, and Ross, 1992; Parkhurst, Shogren, and Bastian, 2004).

Experimental findings presented in the previous chapter point to the important role of the structure of information in coordination games. In the presence of information transmission mechanisms – cheap-talk communication in one experimental condition, and observation of partners' past behavior in the other – players' decisions strongly depend on what the received signals suggest about senders' behavior. Furthermore, these signals also provide some screening of senders' intentions, which facilitates strategic synchronization. On the other hand, the another puzzling

---

<sup>1</sup>The main part of the material presented in this chapter comes from Jacquemet, Luchini, Shogren, and Zylbersztein (2011). Extensions also involve Hanaki, Jacquemet, Luchini, and Zylbersztein (2013).



phenomenon – frequent deviations from the own-payoff-maximizing strategy – occurs robustly and independently of the structure of information in the game.

Herein we explore whether we can create a coordination game environment in which people are committed to match words with actions. According to the social psychology theory of commitment, the acceptance or realization of preparatory acts creates a conceptual bridge between what one thinks and what one does. Research in social psychology suggests that this mechanism may render communication more effective, giving rise to *binding communication* (Joule, Girandola, and Bernard, 2007; Grandjean and Guéguen, 2011). We use a truth-telling oath procedure introduced by Jacquemet, Joule, Luchini, and Shogren (2012) to strengthen the link between players' communications and actions: each player voluntarily signs an oath to tell the truth before he or she enters the lab. Commitment theory and experimental evidence has shown people are more likely to match actions with words when they have freely chosen to commit themselves to doing them through a prior action (see Joule and Beauvois, 1998).<sup>2</sup> Our hypothesis is that the prior action – an oath to tell the truth – will create commitment that will enhance the power of cheap talk communication thereby reducing strategic uncertainty and increasing coordination.

In an experimental design using the classic coordination game of a sender and a receiver (based on Selten, 1975; Rosenthal, 1981),<sup>3</sup> our results are that efficient coordination increases by over 50 percent within the oath treatment. The oath improves efficient coordination by changing the behavior of the sender whose messages become more truthful and focused on efficiency, and whose choices get more efficient. These findings are in line with related economic experiments. People who make promises about future actions after pre-play communication are more likely to keep them (see *e.g.* Ellingsen and Johannesson, 2004; Charness and Dufwenberg, 2006; Vanberg, 2008). The oath also affects the behavior of the receiver. More receivers trust the messages announcing how the sender will play the game. Yet, about half of them still play the unreliable and uncooperative strategy at least once in the experiment. One may thus wonder whether receivers playing  $L$  have a cautious attitude towards the sender or are simply irrational. What we find is that the behavior of receivers under oath is similar to the behavior of subjects playing the same game with automated players and thus facing no uncertainty (Hanaki, Jacquemet, Luchini, and Zylbersztein, 2013), suggesting that remaining coordination failures are due to irrationalities and the oath has no effect on them.

---

<sup>2</sup>In so called low-ball experiments, for instance, subjects are asked their willingness to perform a target behavior before knowing the full costs of the target behavior. Data show that low-balling significantly increases compliance relative to cases in which individuals are asked to perform the target behavior directly. See Cialdini, Bassett, Miller, and Miller (1978) for the seminal experiment, Cialdini and Sagarin (2005) for an overview and Joule and Beauvois (1998) for a comprehensive work on procedures that create commitment.

<sup>3</sup>As discussed in the previous chapter, coordination failures in this context may arise from strategic uncertainty. See also Beard and Beil (1994); Beard, Beil, and Mataga (2001); Goeree and Holt (2001).

Table 2.1: The experimental game

		<b>Player B</b> (Sender)	
		$l$	$r$
<b>Player A</b> (Receiver)	$L$	(9.75; 3)	(9.75; 3)
	$R$	(3; 4.75)	(10; 5)

## 2.2 Description of the experiment

We rely on the normal form game studied in the previous chapter and presented in Table 2.1. The game involves two players: player A, who chooses between actions  $R$  and  $L$ , and player B, who chooses between actions  $r$  and  $l$ . If  $R$  is chosen by player A, player B can maximize both players' payoffs by selecting action  $r$ . Alternatively, player B may choose action  $l$ , which slightly undermines her own payoff but sharply decreases player A's payoff. If, in turn, player A chooses  $L$ , both players' payoffs do not depend on player B's decision – payoffs are the same whatever action is chosen.

Notwithstanding standard game theoretical predictions (discussed thoroughly in Chapter 1), numerous lab experiments (summarized in Table 2.2) report that suboptimal outcomes (that is, other than  $(R, r)$ ) are omnipresent and arise from two kinds of puzzling behaviors. On the one hand, player As are very frequently reluctant to rely on player Bs. On the other hand, a surprisingly large part of player Bs' actions is not oriented towards own-payoff-maximization.

The payoff structure we implement appears as Treatment 1 in the genuine sequential-move game experiment of Beard and Beil (1994). Among several reported payoff schemes, this one induces the most striking behavior among participants: *(i)* the frequency of player As' unreliant choices related to this treatment is remarkable: 65.7%, and *(ii)* this is the only treatment where deviations from the dominant strategy by player Bs were observed (in 17% of all cases where player A made a reliant decision  $R$ ). Chapter 1 of the present thesis reports that roughly 50% of all player As' decisions are unreliant, and that player Bs use dominated strategy in about 20% of all cases once a simultaneous-move game is repeated over multiple periods. Additional treatments moreover show that these results are robust to changes in the payoff structure that eliminate the inequality in players' payoffs in the efficient equilibrium, and enhance the saliency of player Bs' decisions.

**Pre-play communication.** Prior to decision making, player B (hereafter, the sender) transmits a cheap talk signal to player A (hereafter, the receiver), indicating (truthfully or not) her intended decision. We are interested in a fixed-form communication, in which senders choose between three

Table 2.2: Overview of existing experimental evidence

Experiment	Form	Payoff			Outcomes (%)				
		( $L$ )	( $R, r$ )	( $R, l$ )	$L$	$R, r$	$R, l$	$r R$	$r$
Beard, Beil–Tr.1	Seq	(9.75; 3.0)	(10; 5.0)	(3; 4.75)	66	29	6	83	—
Beard, Beil–Tr.3	Seq	(7.00; 3.0)	(10; 5.0)	(3; 4.75)	20	80	0	100	—
Beard, Beil–Tr.4	Seq	(9.75; 3.0)	(10; 5.0)	(3; 3.00)	47	53	0	100	—
Beard et al.–Tr.1	Seq	(1450; 450)	(1500; 750)	(450; 700)	79	18	3	83	—
Beard et al.–Tr.2	Seq	(1050; 450)	(1500; 750)	(450; 700)	50	18	32	64	—
Goeree, Holt–Tr.1	Ext	(80; 50)	(90; 70)	(20; 10)	16	84	0	100	—
Goeree, Holt–Tr.2	Ext	(80; 50)	(90; 70)	(20; 68)	52	36	12	75	—
Goeree, Holt–Tr.3	Ext	(400; 250)	(450; 350)	(100; 348)	80	16	4	80	—
Cooper, Van Huyck–Tr.9	Str	(4; 1)	(6; 5)	(2; 4)	27	—	—	—	86
Cooper, Van Huyck–Tr.9	Ext	(4; 1)	(6; 5)	(2; 4)	21	—	—	—	84
Chapter 1–Baseline (BT1), round 1	Str	(9.75; 3.0)	(3.0; 4.75)	(10; 5.0)	77	23	0	100	80
Chapter 1–Baseline (BT1), rounds 2-10	Str	(9.75; 3.0)	(3.0; 4.75)	(10; 5.0)	48	43	9	84	81
Chapter 1–Baseline (BT1), overall	Str	(9.75; 3.0)	(3.0; 4.75)	(10; 5.0)	51	41	8	84	81
Chapter 1–ET1	Str	(9.75; 5.0)	(5.0; 9.75)	(10; 10.0)	54	33	13	72	73
Chapter 1–ET3	Str	(9.75; 5.5)	(5.5; 8.50)	(10; 10.0)	39	48	13	79	76
Chapter 1–ET4	Str	(8.50; 5.5)	(5.5; 8.50)	(10; 10.0)	25	61	14	82	82
Chapter 1–ET2	Str	(8.50; 8.5)	(6.5; 8.50)	(10; 10.0)	26	70	4	94	94
Chapter 1–BT2	Str	(8.50; 7.0)	(6.5; 7.00)	(10; 8.5)	26	70	4	94	94

**Note.** Several representations of the game have been applied so far, as stated in column 1: simultaneous-move strategic-form game (Str), simultaneous-move extensive-form game (Ext), sequential-move game (Seq). The monetary payoffs displayed in columns 2-4 are in USD in Beard & Beil (1994) and Cooper and Van Huyck (2003), in cents of USD in Goeree & Holt (2001), in Yens in Beard et al. (2001), and in Euros in our experiments.

messages: “I will choose  $r$ ”, “I will choose  $l$ ”, “I will either choose  $l$  or  $r$ ”. The first two messages are informative, while the last one is uninformative in what concerns senders’ intentions.

Cheap talk signal announcing action  $r$  has interesting theoretical properties: namely, it is self-committing, but not self-signalling (Farrell and Rabin, 1996). It is self-committing because, if trusted, it induces the receiver to choose  $R$ , to which the sender’s best response is  $r$  – exactly as the message announces. As argued by Farrell (1988), this is enough to assure signal’s credibility. Yet, player B may still use message “I will choose  $r$ ” to persuade her partner to choose  $R$ , and take any action afterwards. Hence, it is not self-signalling in the sense that it does not reveal sender’s true intentions and, as pointed by Aumann (1990), it should be ignored by the receiver.

Notwithstanding Aumann’s critique, Crawford (1998) suggests that cheap talk communication may nonetheless be used to *reassure* the receiver about sender’s intentions and thus to reduce strategic uncertainty in coordination problems. Subsequent experimental studies (e.g. by Charness (2000)) tend to confirm this point. Ellingsen and Östling (2010) provide a formalization of this idea using the level- $k$  bounded rationality framework. Under the central assumption that agents display a weak preference for truthfulness which is of common knowledge, they demonstrate that

cheap talk messages convey information about sender’s rationality – that is, whether he disregards dominated strategies or not.

In the specific context of the coordination game outlined in Table 2.1, the previous chapter reports that cheap-talk communication helps overcome coordination failure, but is not capable of improving the efficiency of outcomes. Moreover, its performance is similar to another information-transmission mechanism, in which player A observes the entire set of decisions that player B took in interactions with his past partners.

### 2.2.1 Experimental design

**Baseline treatment (No Oath).** Each experimental session consists of 10 rounds of the game presented in Table 2.1. Roles are fixed, so that each participant takes 10 decisions as either receiver or sender. After each interaction, pairs are rematched according to a perfect stranger round-robin procedure that guarantees that each pair of subjects may interact at most once throughout the entire experiment. For the sake of avoiding the end-game effects, we do not reveal to participants the exact number of rounds, but only inform in the experimental instructions that there will be several of them. At the beginning of each round, the sender sends one of the three messages to the receiver by clicking on a relevant button on her computer screen. We explain to the subjects that messages sent by senders do not affect their payoffs, and that they can be followed by any decisions. Once the receiver has confirmed receiving the message, the game moves to the decision-making stage, where the receiver chooses between  $R$  and  $L$ , while the sender – between  $r$  and  $l$ . Instructions inform the participants that decision is first taken by the receiver, then by the sender, and the final payoff depends only on the receiver’s decision should  $L$  has been chosen, and on both partners’ decisions otherwise. At the end of every round, each subject is only informed about her own payoff. Next, individuals are informed that either another round is about to start, or that the experiment ends.

**Oath treatment.** This treatment uses an identical experimental environment as the baseline treatment, except that each subject is asked to sign an explicit oath before entering the lab (see Jacquemet, Joule, Luchini, and Shogren (2012) and Jacquemet, James, Luchini, and Shogren (2010)). The oath is implemented as follows: each subject enters alone and is directed to a monitor at the front of the laboratory. The monitor then offers each subject a form to sign entitled “solemn oath” as presented in Figure 2.1.<sup>4</sup> The *Paris School of Economics* logo on the top of the form and the address at the bottom indicate that it is an official paper; the topic designation and the research number were added so to ensure the credibility. The monitor explicitly points out to the subject before she reads the form that she is free to sign the oath or not, and that participation and earnings are not conditional on signing the oath (subjects are, however, not informed about

---

<sup>4</sup>The word “oath” is written on the form and read by the subject, but never said aloud.

Figure 2.1: Oath form used in the experiment

PARIS SCHOOL OF ECONOMICS  
ÉCOLE D'ÉCONOMIE DE PARIS

**SOLEMN OATH**

Topic: "JZ"; Research number 1842A

I undersigned ..... swear upon my honour that, during  
the whole experiment, I will:

**Tell the truth and always provide honest answers.**

Paris, ..... Signature.....

---

Paris School of Economics, 48 Boulevard Jourdan 75014 Paris - France.

the topic of the experiment when asked to take the oath). The subject reads the form, which asks whether she agrees “to swear upon my honor that, during the whole experiment, I will **tell the truth and always provide honest answers**” (in bold in the original form). Regardless of whether the subject signs the oath, he or she is thanked and invited to enter the lab. The exact wording used by the monitors to offer the oath to respondents was scripted to standardize the phrasing of the oath. One monitor stayed in the lab until all subjects had been presented with the oath, to avoid communication prior to the experiment. Subjects waiting their turn could neither

see nor hear what was happening at the oath-desk.

### 2.2.2 Experimental procedures

We run six experimental sessions (three for each experimental condition), each of them involving 20 subjects – 10 receivers and 10 senders. The no-oath treatment used data from the communication condition discussed in Chapter 1. In the oath treatment, subjects first come one by one to the oath-desk and are exposed to the above described procedure. Although signing the oath was not mandatory, a large majority of subjects accepted to do so. Six subjects did not take an oath. This leads to a 90% acceptance rate. This is in line with previous experiments involving an oath procedure (see Jacquemet, Joule, Luchini, and Shogren, 2012; Jacquemet, James, Luchini, and Shogren, 2010).<sup>5</sup> We apply an intention to treat strategy, so that these six subjects are kept in the statistical analysis – given the number of observations this does not affect the statistical results, though.<sup>6</sup>

In both treatments, participants are randomly assigned to their computers and asked to fill in a small personal questionnaire containing basic questions about their age, gender, education, *etc.* The written instructions are then read aloud. Players are informed that they will play some (unrevealed) number of rounds of the same game, each round with a different partner, and that their own role will not change during the experiment. Before starting, subjects are asked to fill in a quiz assessing their understanding of the game they are about to play. Once the quiz and all remaining questions are answered, the experiment begins. Prior to the first round, players are randomly assigned to their roles – either receiver or sender. Subsequently, they are anonymously and randomly matched to a partner. Then, the sender sends a message to the receiver, after which they are both asked for their choices,  $R$  or  $L$  for receivers, and  $r$  or  $l$  for senders. At the end of every round, each participant is informed solely about her own payoff. Once all pairs complete a round of the game, subjects are informed whether a new round starts. If this is the case, pairs are rematched. Otherwise, a single round is randomly drawn and each player receives the amount in Euros corresponding to her gains in that round, plus a show-up fee equal to 5 Euros.

All sessions took place in the lab of University Paris 1 (LEEP) in between June 2009 and January 2012. The recruitment of subjects has been carried out via LEEP database among individuals who have successfully completed the registration process on Laboratory’s website.<sup>7</sup> The experiment involved a total group of 120 subjects, 72 males and 48 females. 102 of them are students, among which 58 subjects are likely to have some background in game theory due to their

---

<sup>5</sup>This is also a standard acceptance rate for commitment experiments (see Joule and Beauvois, 1998; Burger, 1999).

<sup>6</sup>Note, subjects are never aware of whether their interaction partners signed the oath or not.

<sup>7</sup>The recruitment uses ORSEE (Greiner, 2004); the experiment is computerized through a software developed under REGATE (Zeiliger, 2000).

Table 2.3: Aggregate results

	Round										Overall
	1	2	3	4	5	6	7	8	9	10	
<b>No oath:</b>											
Reliant A ( $R$ )	50.0	36.7	50.0	53.3	66.7	70.0	66.7	70.0	60.0	70.0	<b>59.3</b>
Reliable B ( $r$ )	80.0	70.0	90.0	83.3	66.7	80.0	86.7	83.3	86.7	73.3	<b>80.0</b>
Coordination ( $L, l$ ) or ( $R, r$ )	56.7	46.7	60.0	63.3	66.7	70.0	80.0	66.7	66.7	83.3	<b>66.0</b>
Efficient outcome ( $R, r$ )	43.3	26.7	50.0	50.0	50.0	60.0	66.7	60.0	56.7	63.3	<b>52.7</b>
Miscoordination ( $L, r$ )	36.7	43.3	40.0	33.3	16.7	20.0	20.0	23.3	30.0	10.0	<b>27.3</b>
Miscoordination ( $R, l$ )	6.7	10.0	0.0	3.3	16.7	10.0	0.0	10.0	3.3	6.7	<b>6.7</b>
<b>Oath:</b>											
Reliant A ( $R$ )	70.0	70.0	73.3	83.3	80.0	76.7	76.7	76.7	86.7	80.0	77.3
Reliable B ( $r$ )	96.7	90.0	90.0	90.0	93.3	90.0	100.0	96.7	96.7	90.0	93.3
Coordination ( $L, l$ ) or ( $R, r$ )	73.3	80.0	83.3	86.7	80.0	80.0	76.7	80.0	83.3	70.0	<b>79.3</b>
Efficient outcome ( $R, r$ )	70.0	70.0	73.3	80.0	76.7	73.3	76.7	76.7	83.3	70.0	<b>75.0</b>
Miscoordination ( $L, r$ )	26.7	20.0	16.7	10.0	16.7	16.7	23.3	20.0	13.3	20.0	<b>18.3</b>
Miscoordination ( $R, l$ )	0.0	0.0	0.0	3.3	3.3	3.3	0.0	0.0	3.3	10.0	<b>2.3</b>

**Note.** Columns 1-10 summarize the frequencies of outcomes (defined in rows) as % of all outcomes observed in each round of a given experimental treatment. The last column provides overall results.

field of study.<sup>8</sup> 28% never took part in any economic experiment in LEEP before. Participants' average age is roughly 24. No subject participated in more than one experimental session. Each session lasted about 45 minutes, with an average payoff of approximately 12 Euros.

## 2.3 Results

Table 2.3 summarizes aggregate results by round from both treatments. We see our first key result: the likelihood that players coordinated to the optimal Pareto-efficient outcome increased by nearly 50 percent due to oath – to 75.0 percent optimal coordination with the oath from 52.7% without. Each type of suboptimal coordination –either ( $L, r$ ) or ( $R, l$ )– was less likely in the oath treatment than in the baseline. Accounting for both optimal and suboptimal coordination, the commitment-via-the-oath contributes to coordinating players' actions – the frequency of Nash equilibrium increases to 79.3% in the oath treatment from 66% in the baseline.

**Result 1.** Commitment via the truth-telling oath increased coordination on the socially optimal Nash equilibria to 75 percent of games from 53 percent without the oath.

<sup>8</sup>Disciplines such as economics, engineering, management, political science, psychology, mathematics applied in social science, mathematics, computer science, sociology, biology.

The increase of coordination induced by the oath is explained by significant changes in individual behavior.

Both players' aggregate behavior in the communication treatment is affected by the presence of the oath procedure. For inexperienced subjects (round 1), we observe that oath significantly increases player As' reliance and player Bs' reliability ( $p = 0.051$ ,  $p = 0.094$ , respectively, using one-sided Fisher's exact test). Furthermore, the patterns of both players' behavior remain highly divergent as the game continues: the *lowest* levels of player As' reliance and player Bs' reliability observed throughout the 10 rounds of game (0.700 and 0.900, respectively) in the oath condition happen to coincide with the *highest* levels in the no-oath condition.

Figure 2.2 summarizes both players' decision-making patterns. The degree of receivers' reliance is larger in the oath treatment: out of 10 decisions, an average of 5.9 decisions  $R$  is observed in the baseline treatment, and 7.7 in the oath treatment. A mean difference bootstrap test indicates that the difference is significant ( $p = .029$ ).<sup>9</sup>

Both EDF presented in Figures 2.2.a are similar on the low end, *i.e.* for players who do not play  $R$  often, while discrepancies are visible on the upper end, where the oath treatment induces subjects to play  $R$  more often. We observe that 43.3% of receivers in the oath treatment choose to play  $R$  in all rounds while none of them do so in the baseline treatment (bootstrap proportion test:  $p = .000$ ). The differences between the two EDF are highly significant: the EDF of receivers' behavior in the oath treatment first-order dominates the EDF of receivers' behavior in the baseline ( $p = .003$ ). Substantial differences also appear in senders' behavior, as revealed in Figure 2.2.b. Based on the empirical frequencies of decisions  $r$ , we find that the EDF from the oath treatment first-order dominates the EDF from baseline ( $p = .017$ ). Therefore, in the presence of oath senders are more likely to cooperate with receivers. In particular, 80% of senders choose to play  $r$  in all 10 rounds in the oath treatment, while only 43.3% do so in the baseline. A bootstrap test for equality of proportions indicates that the difference is significant at the 1% level (bootstrap proportion test:  $p = .006$ ).

**Result 2.** The oath shifts sender's actions towards more payoff maximizing decisions.

### 2.3.1 Communication under oath

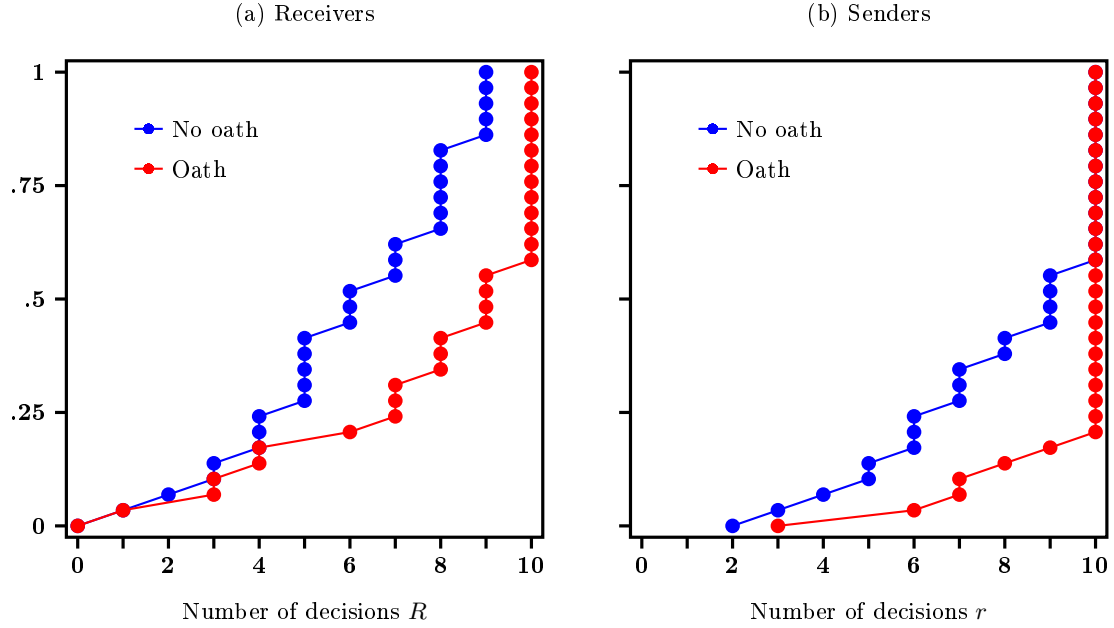
Communication under oath improves coordination, and at the same time shifts players behavior towards more efficient decisions. We now turn to the reasons underlying this changes. First, we look at the patterns of communication by senders. Figure 2.3 shows the empirical distribution functions (EDF) of the number of messages sent by subjects in the role of senders over all 10 decision periods. Each bullet inside the graph indicates an individual, on the  $x$ -axis we present the number of messages he has sent (between 0 and 10), the  $y$ -axis represents the probability of

---

<sup>9</sup>This non-parametric procedure consists of bootstrapping individuals and their corresponding decisions over all ten rounds (999 times), and thus fully accounts for the within-correlation of individual's choices.



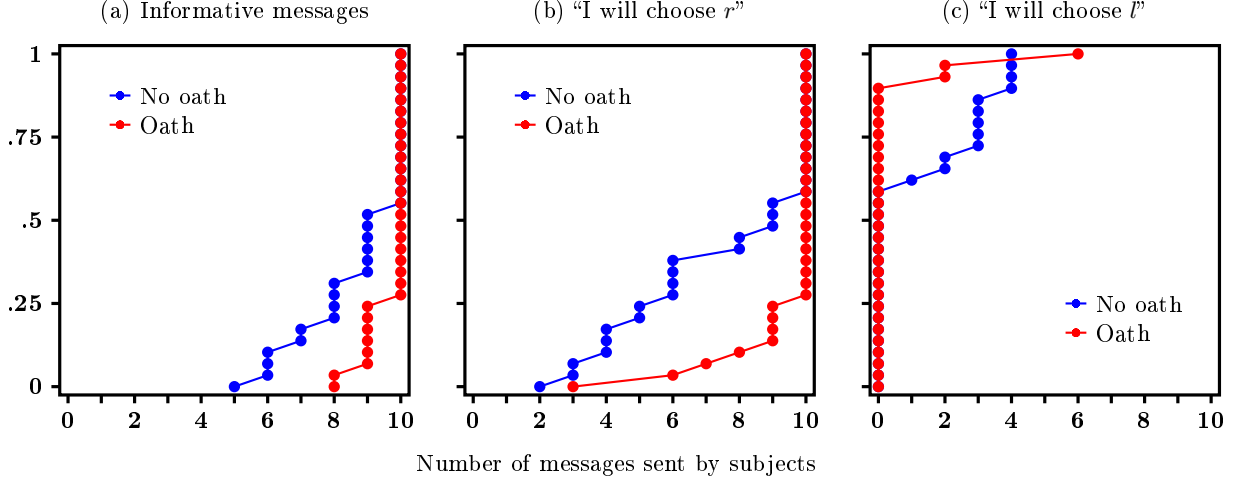
Figure 2.2: Decisions of receivers and senders by treatment



observing an individual who has sent at most a given number of messages. Figure 2.3.a depicts the empirical distribution of informative messages (“I will choose  $r$ ” and “I will choose  $l$ ”) in both treatments. In both cases communication is widely used by senders, since no subject has abstained from sending at least one informative message. The number of messages sent by each subject is relatively high, ranging from 5 messages to 10 messages in the baseline treatment and from 8 to 10 messages in the oath treatment. In particular, 14 subjects out of 30 (46.7%) in the baseline treatment send 10 informative messages, i.e. one such message in every round; 22 subjects out of 30 (73%) do so in the oath treatment. The difference is significant according to a bootstrap test for equality of proportions ( $p = .067$ ). Overall, although senders in the oath treatment seem to use communication more often than senders in the baseline treatment, first-order dominance of the EDF of informative messages in the oath treatment over informative messages in baseline is not statistically significant ( $p = .118$ ).<sup>10</sup> Figure 2.3.b, in turn, presents the analogous EDF exclusively for messages “I will choose  $r$ ”. We find that subjects in the oath treatment display a stronger willingness to signal their credibility than subjects in the baseline treatment: for this class of messages, the EDF from the oath treatment first-order dominates the EDF from the baseline treatment ( $p = .035$ ). Finally, Figure 2.3.c focuses on messages “I will choose  $l$ ”. These messages

<sup>10</sup>This result comes from a bootstrap version of the univariate Kolmogorov-Smirnov test. This modified test provides correct coverage even when the distributions being compared are not entirely continuous and, unlike the traditional Kolmogorov-Smirnov test, allows for ties (see Abadie, 2002; Sekhon, 2011).

Figure 2.3: Communication behavior of senders by treatment



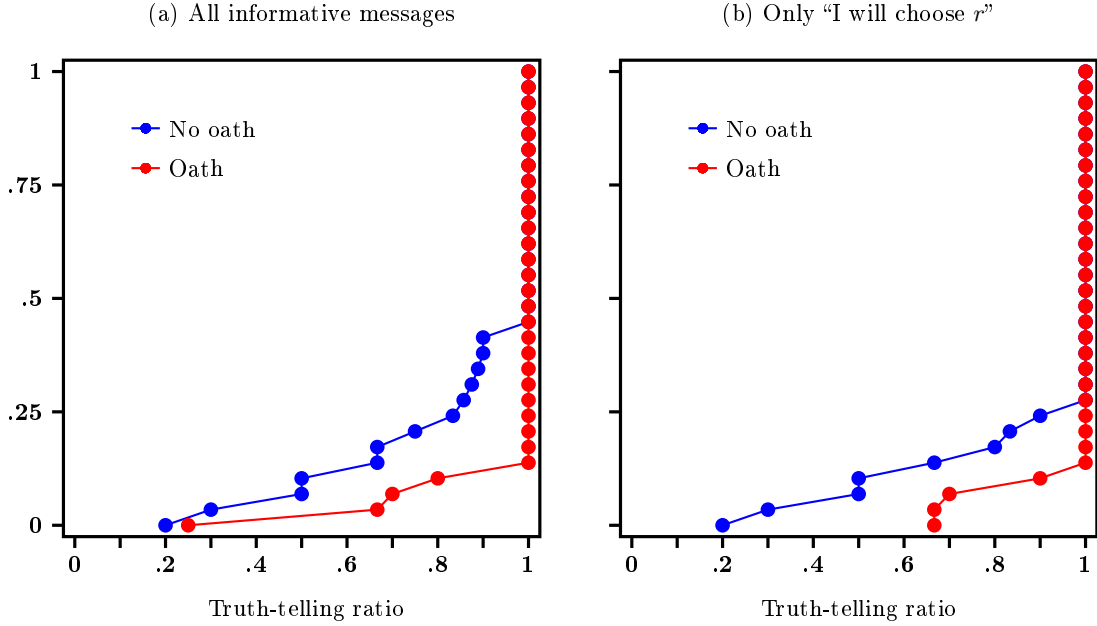
are seldom used – only 3 subjects out of 30 send them at least once in the oath treatment, whereas 12 of them out of 30 do so in the baseline; this difference is significant at the 5% level ( $p = .012$ ). Altogether, we find that the individual communication behavior does not vary much in quantity – since senders use communication to a similar extent in both treatments – but does change in quality – in the oath treatment messages “I will choose  $r$ ” are sent significantly more often, while messages “I will choose  $l$ ” are sent significantly less often.

**Result 3.** The oath both (1) increases the informativeness of transmitted signals, and (2) changes the signal’s structure by substantially increasing the use of reassuring announcements.

We now explore the link between messages and actions through the truth-telling behavior of senders. For sender we calculate the proportion of cases in which the action is coherent with the message (given it is informative) within all the cases in which an informative message is sent. We call this the truth-telling ratio; Figure 2.4.a presents the EDF from both treatments. Looking at the figure, we find striking evidence that misinforming one’s partner about intended move is substantially more widespread without the oath. First, 26 of 30 senders (87%) always reveal their actual intentions when sending an informative message in the oath treatment relative to only 17 of 30 player Bs (57%) in the baseline treatment. The difference is statistically significant with  $p = 0.016$  according to a bootstrap proportion test. Second, the EDF from the oath treatment significantly first-order dominates the EDF from baseline ( $p = 0.067$ ). In Figure 2.4.b, we represent the same truth-telling ratio using only messages “I will choose  $r$ ”. Here the difference between the two treatments is more ambiguous. The EDF of the truth-telling ratio from the oath treatment still first-order dominates the EDF from baseline, but the result is not statistically significant ( $p = 0.586$ ).

**Result 4.** The oath improves the truthfulness of announcements.

Figure 2.4: Truthfulness of senders by treatment (Empirical distribution function)



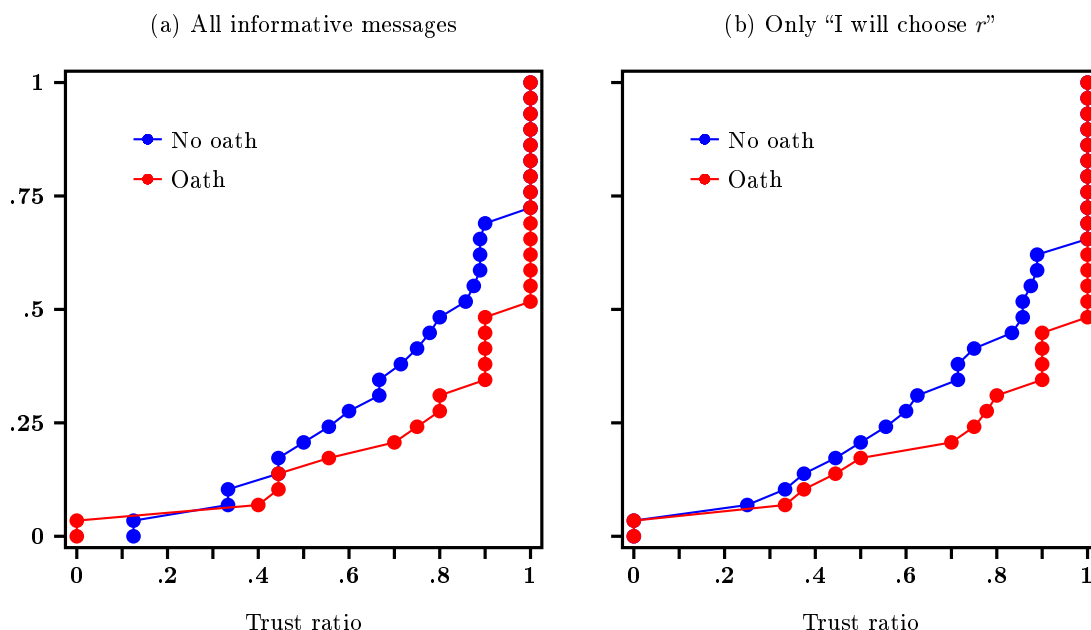
Last, in Figure 2.5 we look at receivers' perception of information obtained from their partners, conditional on the presence of oath. More precisely, for each individual in the role of the receiver we calculate the proportion of cases in which he gives the best response to an informative signal (that is, when the receiver plays  $R$  should the signal from the sender announce action  $r$ , or receiver's move  $L$  follows message announcing move  $l$  by sender) among all the cases when the received message happens to be informative. We call this proportion the trust ratio, and present its EDF in Figures 2.5.a and 2.5.b. The first figure suggests that oath has an influence on how receivers react to messages: the EDF obtained in the oath treatment first-order dominates the EDF obtained in the baseline ( $p = .035$ ). This remains true when restricting attention to messages "I will choose  $r$ ", as it can be seen in the second figure. The differences in behaviors are less pronounced but still significant ( $p = .067$ ).

**Result 5.** The oath improves the receiver's trust toward the announcement.

## 2.4 Commitment without communication

We observe two simultaneous shifts in behavior in our previous treatments. First, in the oath condition both players behave more rationally: both player As and player Bs take decisions leading to the efficient outcome more often. Second, coordination improves because communication is more efficient: player Bs announce decision  $r$  more often and act more truthfully ; and this announcement

Figure 2.5: Messages trusted by player As (Empirical distribution function)



is trusted and followed by player As. This is enough to conclude that communication under oath improves efficiency, but this is not enough to conclude why this is so: coordination can result from either a more efficient communication technology applied to otherwise unchanged behavior, or to the reverse – improved behavior implemented through an unchanged communication technology – or both. To disentangle between these possible explanations, we design a robustness experiment to elicit the direct effect of the oath on behavior – without communication.

### 2.4.1 Experimental design

In order to evaluate the specific effect of oath on behavior, we introduce one change to our previous design: this time, player Bs cannot announce their intention to player As. In the **baseline-no communication** condition, subjects are randomly allocated permanently to the role of either player A or player B for ten rounds of the simultaneous-move game presented in Table 2.1. Pairs are rematched from one period to the other using a perfect stranger, round-robin matching procedure, and subjects earnings from the experiment are computed as the sum of a 5€ show up fee and the result of their payoff in a single, randomly drawn round. In the **oath-no communication** treatment, subjects are offered to sign a truth-telling oath before entering the lab to participate in the experiment.

For each of the two conditions, our data come from 3 sessions, involving 20 subjects each. The data for the baseline-no communication condition are taken from the baseline treatment 1

Table 2.4: Coordination under oath, without communication: Aggregate results

	Round										Overall
	1	2	3	4	5	6	7	8	9	10	
No oath:											
Reliant A ( $R$ )	23.3	36.7	50.0	50.0	60.0	60.0	56.7	53.3	56.7	43.3	49.0
Reliable B ( $r$ )	80.0	80.0	90.0	83.3	73.3	80.0	76.7	76.7	80.0	86.7	80.7
Coordination ( $L, l$ ) or ( $R, r$ )	43.3	43.3	53.3	60.0	66.7	60.0	60.0	43.3	63.3	36.7	53.0
Efficient outcome ( $R, r$ )	23.3	30.0	46.7	46.7	50.0	50.0	46.7	36.7	50.0	33.3	41.3
Miscoordination ( $L, r$ )	56.7	50.0	43.3	36.7	23.3	30.0	30.0	40.0	30.0	53.3	39.3
Miscoordination ( $R, l$ )	0.0	6.7	3.3	3.3	10.0	10.0	10.0	16.7	6.7	10.0	7.7
Oath:											
Reliant A ( $R$ )	33.3	40.0	40.0	40.0	46.7	46.7	43.3	46.7	40.0	50.0	42.7
Reliable B ( $r$ )	86.7	66.7	83.3	93.3	80.0	83.3	86.7	90.0	80.0	90.0	84.0
Coordination ( $L, l$ ) or ( $R, r$ )	46.7	46.7	36.7	40.0	60.0	43.3	43.3	43.3	40.0	46.7	44.7
Efficient outcome ( $R, r$ )	33.3	26.7	30.0	36.7	43.3	36.7	36.7	40.0	30.0	43.3	35.7
Miscoordination ( $L, r$ )	53.3	40.0	53.3	56.7	36.7	46.7	50.0	50.0	50.0	46.7	48.3
Miscoordination ( $R, l$ )	0.0	13.3	10.0	3.3	3.3	10.0	6.7	6.7	10.0	6.7	7.0

**Note.** Columns 1-10 summarize the frequencies of outcomes (defined in rows) as % of all outcomes observed in each round of a given experimental treatment. The last column provides overall results.

(BT1) from the previous chapter. The oath-no communication sessions have been run in October 2012, with all subjects but two (58/60) freely deciding to sign the oath. Among the total of 120 participants (69 males and 51 females), 105 are students.<sup>11</sup> Subjects' average age is 23, 59% took part in an experiment before. The average payoff is approximately 12 Euros including the show-up fee of 5 Euros.

## 2.4.2 Results

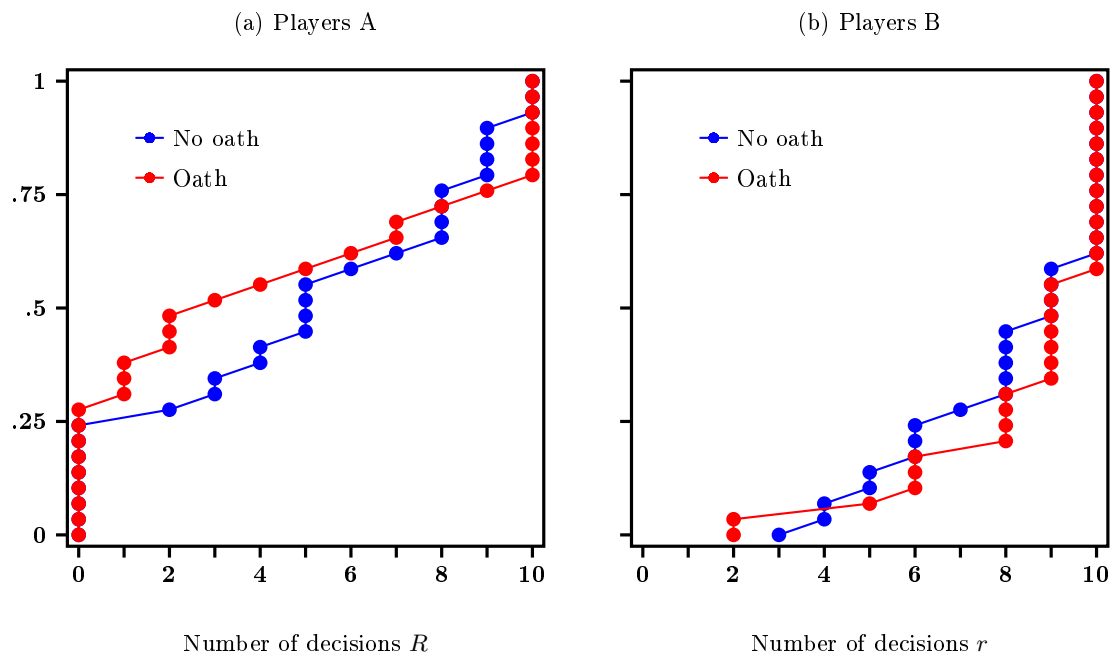
Table 2.4 summarizes aggregate results by round from the no-communication treatments. We observe that the oath does not induce differences in aggregate behavior.

Unlike the communication condition, in the absence of communication one does not observe that oath significantly increases the frequencies of actions  $R$  or  $r$  among unexperienced players ( $p = 0.284$ ,  $p = 0.365$ , respectively, using one-sided Fisher's exact test for round 1 data). Then, the evolution of behavior and its outcomes is very similar for both treatments.

In particular, the likelihood that players coordinate on the efficient outcome is 41.3% in the baseline-no communication and 35.7% in the oath-no communication treatment. Coordination, whereas it is optimal or not, accounts for 53.0% of the outcomes in baseline and 44.7% when

<sup>11</sup>With 54 students enrolled in programs in economics, engineering, management, political science, psychology, mathematics applied in social science, mathematics, computer science, sociology, biology.

Figure 2.6: Observed decisions without communication



subjects took an oath prior to the experiment.

Figure 2.6 summarizes the decision-making pattern of player As and player Bs. Individual data point to the same conclusion: the oath does not induce any change in behavior when communication is not allowed. Out of 10 decisions, 4.9 decisions  $R$  are on average taken by player As without oath, as opposed to 4.3 with an oath – the difference is not significant ( $p = .535$ ). The average number of decisions  $r$  from player Bs equals 8.1 without oath and 8.4 with an oath – the difference is not significant,  $p = .564$ . Bootstrap distribution tests indicate that the EDF are not significantly different with  $p = .416$  for player As and  $p = .676$  for player Bs. These results allow us to state our final result 2.4.2:

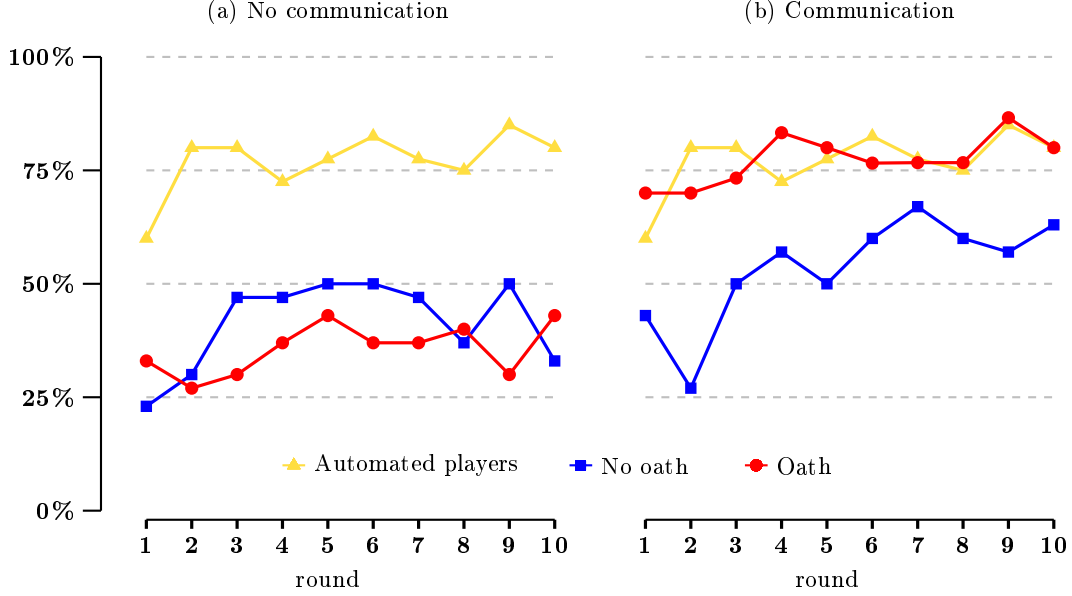
**Result 6.** The oath alone does not improve efficient coordination, but it can show its effects through communication.

### 2.4.3 Communication under oath and strategic uncertainty

Communication under oath increases efficiency by 50% and subjects coordinate on the Pareto efficient outcome 75% of the times. Although this is a significant improvement in efficiency, the open question is whether one could further reduce coordination failures through an even more efficient communication technology.

To provide evidence on that issue, we rely on the results of Hanaki, Jacquemet, Luchini,

Figure 2.7: Efficient coordination in human subjects and automated players treatments



and Zylbersztejn (2013) based on a game without communication identical to our **baseline-no communication** treatment but with one important exception: player As face automated players instead of human subjects in the role of player Bs. Automated players are programmed so as to always play the payoff-maximizing strategy  $r$ , and this is common knowledge to the subjects in the experiment.<sup>12</sup> Apart from this change, the whole experimental design is the same as ours. Note player As now have no “rational” reason to use action  $L$  as the automated players treatment totally removes strategic uncertainty.<sup>13</sup>

In Figure 2.7, we confront the likelihoods of attaining the Pareto-efficient outcome in the automated treatment and our four human treatments: no-communication treatments in Figure 2.7.a and communication treatments in Figure 2.7.b. At the aggregate level, efficient coordination

<sup>12</sup>A brief summary of this experiment – its procedures and main results – is provided as a supplementary material.

<sup>13</sup>A growing body of literature uses robots to control for subjects’ beliefs about the behavior of their opponents. Robots may not follow the equilibrium strategies, however, because the focus of these papers is not on disentangling the effects of strategic uncertainty and bounded rationality. For instance, Cason and Friedman (1997) consider robot players in an experiment on price formation in a simple market institution. They mainly concentrate on robot traders who follow the Bayesian Nash Equilibrium strategy (BNE robot) but also introduce a “Revealing Robot” to investigate whether the convergence towards the equilibrium is due to human subjects mimicking the behavior of BNE robots or is their best response to BNE robots. Their results support the latter hypothesis. In other studies, robots are used to replicate the observed behaviors of subjects in the past (e.g., Ivanov, Levin, and Niederle, 2010) or to make some players follow the predetermined distribution of boundedly rational behaviors (e.g., Embrey, Fréchette, and Lehrer, 2012).

amounts to 77% of observed outcomes in the automated players treatment and is relatively stable across all rounds after round 1. Efficient coordination without communication, with or without oath, is always far below that of the automated players treatment, reaching a maximum of merely 50% in the no oath condition. When communication is allowed but no oath taken, efficient coordination is still below the level obtained in the automated players treatment. It is only when subjects communicate under oath that efficient coordination attains the level observed in the automated benchmark.

To further explore this comparison at the individual level, Figure 2.8 reports the EDF of the number of decisions  $R$  taken by each player A throughout the experiment in the automated players treatment together with the EDF observed in our four treatments. As shown in panel (a) the EDF from the automated players treatment first order dominates that of **baseline-no communication** ( $p = .001$ ) and that of **oath-no communication** ( $p = .004$ ). In Figure 2.7.b, the EDF of the automated players treatment is plotted with that of our communication treatments. The EDF of decisions in the **baseline-communication** condition is significantly dominated by the EDF in the automated players treatment ( $p < .001$ ). EDF from the **oath-communication** condition and in the automated players treatment are not significantly different ( $p = 0.988$ ).

To summarize, player As facing automated players and player As exposed to communication under oath exhibit exactly the same pattern of behavior. By construction, strategic uncertainty plays no role in Player As choosing  $R$  against automated players: this comparison hence suggests that committed communication through the oath manages to resolve all failures to reach the efficient outcome which are due to strategic uncertainty.

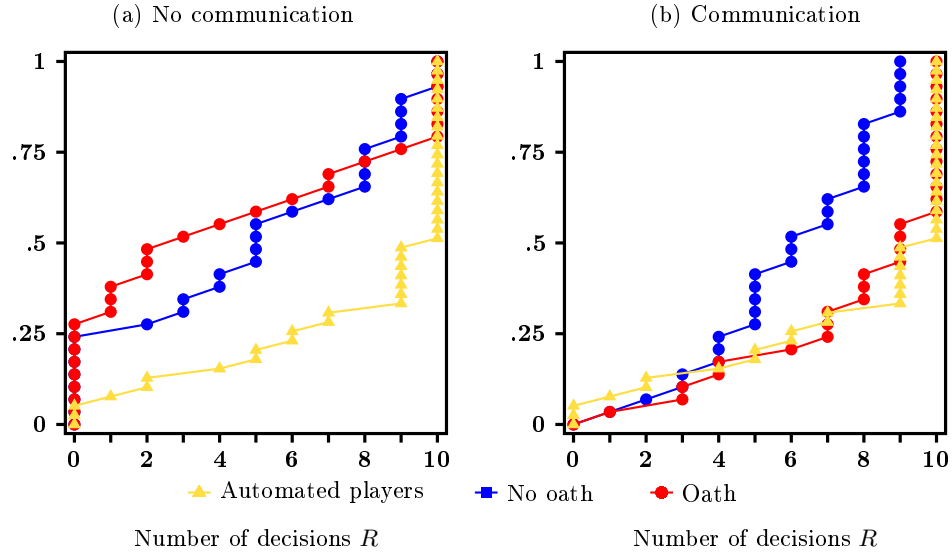
## 2.5 Conclusion

Overcoming coordination failure through economic design can be enhanced with more insight into non-monetary commitment devices. If people can freely choose their coordination partner, they can choose people they already know are trustworthy, so commitment is guaranteed and coordination is efficient (see Riedl, Rohde, and Strobel, 2011). But in many cases, people must coordinate actions with people they do not know. An institution often prescribed to encourage efficient coordination – communication – may not eradicate coordination failure if agents' commitment to backing words with deeds is unclear. This paper investigates whether the social psychological idea of commitment via a truth-telling oath can improve the degree to which communication reduces strategic uncertainty, thereby increasing coordination on efficient outcomes. We create commitment by having our players sign a voluntary oath to tell the truth prior to entering the laboratory (Jacquemet, Joule, Luchini, and Shogren, 2012). Our results suggest that the truth-telling oath significantly increased optimal coordination by nearly 50 percent.

The oath has a significant effect on behavior. In our communication game, senders are more likely to send informative and reassuring messages, and then back up their words with actions.



Figure 2.8: Player As' behavior against automated players



Receivers, in turn, are less prone to change their behavior. More receivers trusted the messages received under oath than without it, but still half of them remained wary. The oath appears to create commitment in senders' signals/actions and in receivers' responses, but it is the oath embedded in communication that creates commitment, and not the oath alone. Furthermore, improved communication did not translate into 100% efficiency in the coordination game, but the level of efficiency is not different from the behavioral equilibrium observed in the game where strategic uncertainty is ruled out as a decision-making motive. This suggests that coordination under oath removes strategic uncertainty, but has no effect on bounded rationality.

## Supplementary material

### Experiment by Hanaki, Jacquemet, Luchini, and Zylbersztejn (2013): a summary

#### Experimental design and procedures

This experiment uses a  $2 \times 2$  design, based on two different payoff matrices – Baseline Game, or BG, which corresponds to BT1 from chapter 1, and Egalitarian Game, or EG, which corresponds to ET2 in chapter 1 (both are presented in Table 2.5), and two different interaction mechanisms. The Human treatment relies on experiments presented earlier in which human subjects in the role of player A are matched with other human subjects in the role of player B, playing 10 periods of the game within changing pairs. In the Robot treatment, human subjects taking the role of player A are matched with a computer program taking the role of player B. The computer is preset to always take decision  $r$  in the game. Subjects in this treatment are clearly informed both about the behavior of the computer program – “**the computer will choose  $r$  in every round, with no exception**” (bold in the original instruction sheet) – and about the fact that subjects are matched with the computer program. This is the only difference in rules and procedures between Human and Robot treatments. 40 (38) participants took part in Robot-BG (Robot-EG) experiments (acting solely as player As), while 120 subjects (30 player As and 30 player Bs per game) participated in Human-BG and Human-EG experiments, all of which have been carried out in Paris.

Table 2.5: The experimental games

player B			player B		
player A	$l$	$r$	player A	$l$	$r$
$L$	(9.75;3)	(9.75;3)	$L$	(8.5;8.5)	(8.5;8.5)
$R$	(3;4.75)	(10;5)	$R$	(6.5;8.5)	(10;10)

(a) Baseline Game

(b) Egalitarian Game

#### 2.5.1 Bounded rationality and strategic uncertainty

Figure 2.9 shows the dynamics of the frequencies of  $R$  decisions for both games and both treatments. It indicates a substantial increase in the proportion of  $R$  decisions when subjects interact with a computer rather than human subjects: in round 1, the percentage of player As who choose

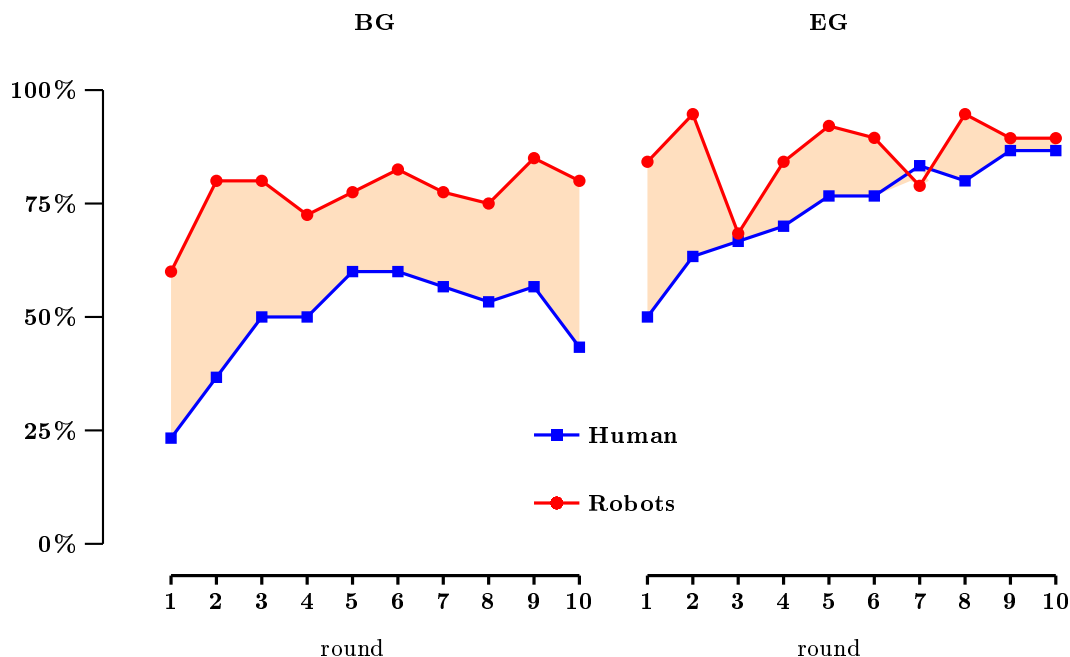


Figure 2.9: Proportion of  $R$  decisions across rounds and treatments

$R$  increases from 23% to 60% in BG and from 50% to 84.2% in EG.<sup>14</sup> Also, the overall percentage of  $R$  decisions increases from 49% to 77.0% in BG, from 77.0% to 86.6% in EG.<sup>15</sup> A bootstrap proportion test indicates that this increase in proportion of  $R$  choices is significant at a 1% threshold in BG ( $p = .001$ ) and at a 10% threshold in EG ( $p = .074$ ). It is also clear from Figure 2.9 that eliminating strategic uncertainty does not lead all player As to select the own-payoff-maximizing choice  $R$  in all 10 rounds. On average, we observe that 23% (13.4%) of choices are  $L$  in BG (EG) when interacting with the computer – even though there is no strategic uncertainty. Such  $L$  choices in Robot treatments can only be interpreted in terms of the bounded rationality of subjects.

Using Robot treatments as a baseline for bounded rationality, we can now estimate to what extent strategic uncertainty explains the choice of the secure option  $L$  in Human treatments. As indicated above, we observe that 51% of choices are  $L$  in human BG and 23% in human EG. Subtracting the proportion of  $L$  choices in Robot treatments from the proportion of  $L$  choices in Human treatments gives us how much of the  $L$  choices can be explained by strategic uncertainty:

<sup>14</sup>The null hypothesis that the observed behavior in round 1 is the same for both Human and Robot treatments is rejected with  $p$ -value 0.0033 for BG and 0.0035 for EG (Fisher's exact test, two-tailed).

<sup>15</sup>For Human treatments, a low degree of reliance from player As (only 49% of actions  $R$  throughout the game) happens to coincide with a high likelihood of inefficient behavior from player Bs (only 81% of decisions  $r$  in the game), which may suggest that player As' behavior stems from strategic adaptation to this risk. However, even when player Bs' behavior improves dramatically (attaining 94% of actions  $r$  in EG), player As remain unreliable very often – in 23% of all cases, which suggests that this behavior is not solely a matter of player As' reaction to the strategic context of the game.

28% (51.0% - 23%) in BG and 9.6% (23% - 13.4%) in EG.

Thus, the key result of this paper is that strategic uncertainty accounts for about half of the observed failures to achieve Pareto efficient outcome in these two simple games. The other half can only be interpreted as a result of (widely understood) individual bounded rationality, which suggests that some subjects do make non-strategic decisions.

## Chapter 3

# Strategic signaling or emotional sanctioning? An experimental study of ex post communication in a repeated public goods game

### 3.1 Introduction

A growing body of experimental studies find that institutions may promote pro-social norms of behavior in economic contexts where individual rationality conflicts with social interest. These institutions are predominantly based on two mechanisms: *ex ante* communication and *ex post* monetary punishment.<sup>1</sup> This paper is devoted to an institution that has been only recently brought to economists' attention as a means of inducing efficiency in social dilemmas: *ex post* communication. Although empirical results are encouraging, the root underlying the effect of this institution, especially in a repeated interaction, is still obscure. This study provides a novel empirical testbed for two mechanisms by which *ex post* communication may potentially affect behavior in a repeated interaction: one is related to signaling, the other suggests that agents' utility depends on only from the monetary outcome of their decisions, but also from other agents' perception (*i.e.* approval or disapproval) of these decisions.

In a seminal contribution, Masclet, Noussair, Tucker, and Villeval (2003) use a repeated four-person public goods game, based on the voluntary contribution mechanism (VCM in short). After each round, every subject observes his group members' contributions, sends a message containing

---

<sup>1</sup>See Masclet, Noussair, and Villeval (forthcoming) for an extensive review of this substantial literature. A notable exception is the experiment by Wilson and Sell (1997) who report no effect of communication on contributions to the public good.

*disapproval points* to each group member, and is informed about the sum of points received from others. Masclet et al. conjecture that *ex post* communication in a repeated interaction may affect subjects' behavior in two ways. First, it may serve as an information transmission device prior to the next round – for instance, signal may convey a warning that the sender will decrease his future contribution unless the receiver increases his. Second, being aware of others' opinions may affect emotions – for instance, people may display an aversion to being disapproved/a preference for being approved, and act less opportunistically as to avoid/deserve it (Holländer (1990)).<sup>2</sup> They argue that these two effects may be separated by confronting subjects' behavior under partner and stranger matching. Intuitively, in the former the effect of *ex post* communication may stem from both strategic information transmission and disapproval-aversion, while in the latter behavioral effects may only be a matter of subjects' aversion to disapproval. In their experiment, partner matching yields significantly higher contributions than stranger matching, which supports the information transmission hypothesis. In a related study, Peeters and Vorsatz (forthcoming) use a similar experimental game and introduce treatments in which every subject may transmit an emoticon (a frowny in one conditions, a smiley in the other) to each partner, and then is informed about the number of emoticons he received. Confronting patterns of behavior under partner matching (where a moderate treatment effect is observed) and stranger matching (where there is virtually no treatment effect), the paper concludes that *ex post* messages are unlikely to involve emotions, but rather facilitate the exchange of information before upcoming round.

On the other hand, an important body of experimental studies suggest that the effect of *ex post* communication in social dilemmas is driven by emotions. López-Pérez and Vorsatz (2010) report that the availability of fixed-form, post-play messages containing judgments of other participants' decisions makes subjects more cooperative in a one-shot prisoner's dilemma game. Ellingsen and Johannesson (2008) and Xiao and Houser (2009) identify the same effect on altruism in subjects playing a dictator game when a free-form, post-play communication is possible. A study by Czap, Czap, Khachatryan, Burbach, and Lynne (2011) implements a two-stage game in which a common-pool resource is used by a group of subjects, out of which some have private incentives to produce publicly undesirable externalities. They find that the reception of a negative emotional feedback (via frownies) after the first stage reduces externalities in the second stage, while providing positive feedback (via smilies) is detrimental with this respect. Dugar (2010) studies a repeated weakest-link coordination game and finds that the introduction of disapproval points helps groups of players converge towards the Pareto-superior Nash equilibrium, while approval points bring no improvement with this respect, not inhibiting a fast convergence to the Pareto-inferior Nash

---

<sup>2</sup>Masclet, Noussair, Tucker, and Villeval (2003) refer to *ex post* communication as *non-monetary sanctioning*, provided that in a companion treatment attributed points are transformed into monetary punishment. Similar nomenclature (like *informal sanctions* or *non-monetary punishment*) is also used in other papers discussed in this section. Since my goal is to test whether this institution is actually used to inform or to sanction people, I consistently use a more neutral term – *ex post* communication.

equilibrium. Dugar (2013) finds that the availability of disapproval points generates higher contributions than approval points in a fixed-group VCM game based on Masclet, Noussair, Tucker, and Villeval (2003), and that combining *both* kinds of points brings a further improvement with this respect. Dugar attributes the effects observed in both of these studies to subjects' asymmetrical fragility to negative and positive emotional feedback.

However, a methodology based on a direct comparison of outcomes arising under partner and stranger matching may be misleading. An extensive survey by Andreoni and Croson (2008) not only reveals that both matching schemes (partner and stranger) are very likely to affect *per se* subjects' behavior in repeated public goods games, but also that the relationship between both protocols is ambiguous – in some experiments partner matching provides higher contributions, other studies report the opposite. In general, comparing outcomes observed under stranger and partner matching dissembles the issue of disentangling proper treatment effects from simultaneous variation in behavior due to applied matching method.

Herein, I propose a novel test of the hypotheses put forward by Masclet, Noussair, Tucker, and Villeval (2003). The design of my experiment builds on Masclet et al.'s sound intuition that under both partner and stranger matching *ex post* communication may affect subjects' emotions, but only the former also allows for strategic signaling. At the same time, the experiment neutralizes unwanted effects of matching protocol on contributions.

### 3.2 Empirical strategy

The experimental methodology introduced by Masclet, Noussair, Tucker, and Villeval (2003) and subsequently used by Peeters and Vorsatz (forthcoming) has one important virtue and one important inconvenience when it comes to disentangling the signaling dimension and the emotional dimension of *ex post* communication in a repeated interaction. Its advantage is that it identifies environments where both of these phenomena either can or cannot coexist. The main disadvantage is that these environments may *per se* affect behavior simultaneously to the content of messages, and that their effects are highly unpredictable.

As a simple illustration of this problem, consider a situation where only two factors affect contributions: *ex post* communication and matching protocol. Since both effects arise simultaneously, future contributions are affected not only by past communication, but also – unobservedly – by the matching protocol. The two effects should not be considered as orthogonal: assuming that the content of communication is a reaction to experienced behavior (which is the case in most studies of *ex post* communication, including this one), future messages depend on past contributions, so that the very nature of communication is also (indirectly) influenced by the matching protocol. As a consequence, the presence of a hidden matching protocol effect in a repeated game aggravates the isolation of the actual relationship between communication and contributions. Masclet, Noussair, Tucker, and Villeval (2003) provide a way to control for matching protocol effects by combining

within- and between-subject approach: in every session subjects play a sequence of rounds without communication, which is followed by an analogous sequence with communication, and then another sequence without communication, holding the matching protocol constant. They argue that since subjects' behavior is similar in the initial sequence under both matching schemes, thus the differences observed in the second stage are unlikely to stem from matching protocol effects. Although their claim is absolutely fair, it should be also noticed that a conclusive comparison in the second stage is only possible in the absence of a systematic partner-stranger difference in the first stage – otherwise, matching protocol effects observed in the first stage would uncontrollably interact with the effects of communication in the second stage. Moreover, a scenario where observations are neutral to matching protocol effects is far from being a regularity in lab experiments on public goods games.<sup>3</sup> Peeters and Vorsatz (forthcoming) use a classical between-subject design and observe an important scope of matching protocol effects: in each round of every treatment, partner matching induces higher contributions than stranger matching (see the working paper version of their study, Peeters and Vorsatz (2009)). The most general solution to this issue boils down to introducing an experimental protocol where actions are uncorrelated with matching scheme.

The main challenge underlying my empirical strategy is to create an experimental environment that allows to differentiate the features of *ex post* communication, at the same time ruling out the unobserved effect of matching protocol on decisions. To this end, I introduce an innovative uniform matching protocol. In each round of a repeated game, subjects decide upon their level on contribution before learning whether their group prevails until next round. Consequently, I neutralize the undesired forward-looking strategic effects of matching protocol on contributions, while controlling for the backward-looking factors (as discussed in the next section). *Ex post* communication, in turn, only takes place after the fate of groups is known to subjects. Consequently, the effect of communication can be captured in two different strategic contexts: when groups prevail from one period to another, and when they change between rounds. In line with the Masclet et al. original argument, strategic information transmission is solely possible in the from case, while referring to emotions may occur in both cases.

Like Masclet, Noussair, Tucker, and Villeval (2003) and Peeters and Vorsatz (forthcoming), I implement non-verbal communication. In order to assure the interpretability of messages, my experimental setting includes the following characteristics. First, the VCM game is played by groups of two subjects, so in each round every participant learns about other group member's contribution, sends a message and receives one in return. Consequently, messages may be easily matched to actions, which creates an environment that (i) facilitates agents' comprehension of non-verbal content, and (ii) allows the experimenter to establish a relationship between individual

---

<sup>3</sup>For instance, these effects are absent in only 4 out of 13 experiments overviewed by Andreoni and Croson (2008). In 5 cases partner matching brings higher contribution than stranger matching, in the remaining 4 cases the opposite occurs.



messages and individual contributions.<sup>4</sup> Second, it is of common knowledge that no group ever re-appears after having been re-matched, which rules out strategic information transmission between subjects who are about to cease interacting.<sup>5</sup>

### 3.2.1 Experimental game and conditions

**Experimental game.** Pairs of subjects play the following version of VCM game. Each player holds an initial endowment of 15 Experimental Currency Units (ECU), and may contribute any part of it to the common pool.<sup>6</sup> Decisions are made simultaneously and the amount accumulated in the common pool is then multiplied by 1.5 and re-transferred to group members in equal parts. Thus, the gain of player  $i$  who contributes  $N_i$  and interacts with player  $j$  who contributes  $M_j$  equals:

$$Gain_i = 15 - 0.25 \times N_i + 0.75 \times M_j \quad (3.1)$$

Although social welfare is maximized when each player contributes his entire endowment, the dominant strategy is to contribute nothing.<sup>7</sup> Therefore, the standard game theory predicts that in the unique Nash equilibrium all players contribute nothing.

**Baseline condition.** In the baseline condition (BC for short), the game is repeated in the following way. In each occurrence, subjects *(i)* decide upon their contribution to the common pool, *(ii)* learn whether their current group prevails until next round, and finally *(iii)* observe other group member's contribution, as well as personal gain. Subjects are informed that in each round their groups survive with a 50% chance, that this process is purely random, and that every change is permanent – groups that disappear cannot re-appear in the future. In all rounds following round 1, an announcement prior to stage *(i)* reminds subjects whether their group has changed

---

<sup>4</sup>It should be mentioned that studies by Masclet, Noussair, Tucker, and Villeval (2003) and Peeters and Vorsatz (forthcoming) (that use a four-person VCM game) explore the formation and the role of monetary and non-monetary mechanisms of peer pressure, which differs substantially from my objective. In these experiments, players evaluate the actions of their group members, and receive a bundled message containing the sum of three evaluations attributed to him by others (via either points or emoticons), which makes the link between a particular message and a particular action much more ambiguous.

<sup>5</sup>In contrast, the stranger matching procedures adopted by Masclet, Noussair, Tucker, and Villeval (2003) and Peeters and Vorsatz (forthcoming) are not "leakproof": the same subjects may interact multiple times throughout the experiment, even though they can never ascertain each other's identity. Masclet, Noussair, Tucker, and Villeval (2003) inform subjects that the probability of being matched with anyone for two consecutive rounds is null, and that the same applies to being part of the same group of four players more than once throughout the experiment. Peeters and Vorsatz (forthcoming), in turn, allow for a completely random matching. Clearly, this makes strategic information transmission between subjects across rounds much more difficult, but not impossible.

<sup>6</sup>In the lab implementation, contributions may only take integer values between 0 ECU and 15 ECU.

<sup>7</sup>To avoid framing effects, instructions use neutral phrasing: I use expressions such as *players* and *group members* rather than *partners*, and *contributions* are never related to *cooperation*. See Rege and Telle (2004) for evidence on the effect of problem framing in public goods experiments.

with respect to the previous period. The important issue of the asymmetry of information about players' past behavior between maintained and newly formed groups is addressed in the following way. Before stage *(i)*, members of newly matched pairs are informed about the decision that was taken recently by their current group member in his former group.<sup>8</sup> Although the amount of historical information may vary between pairs due to the random nature of the re-matching protocol, this design maintains the *minimum* level of knowledge about group members' past.<sup>9</sup>

**Evaluation condition.** Evaluation condition (EC in short) encompasses the three stages forming BC, as well as the current-group-status reminder. In addition, in stage *(iv)* subjects are asked to express their opinion about group member's decision by attributing him a certain number of points (between 0 and 10). Experimental instructions state that *a high number of points expresses disapproval: 10 points correspond to the strongest disapproval, while 0 points correspond to the weakest disapproval*, and that *attributed points do not affect either participant's gains for the experiment*.<sup>10</sup> Then, each subject is informed about the number of points he received from the other group member. If groups change between periods  $t - 1$  and  $t$ , prior to stage *(i)* subjects not only learn about the decision taken by their current group member in  $t - 1$  (like in BC), but also about the number of points he received. The provision of this additional information stems from the emotional-based hypothesis: if subjects are sensitive to received points for emotional reasons, then the evaluation of their past actions may *pre se* affect their future action, regardless of whether re-matching occurred in the meantime or not. Therefore, if the emotion-based hypothesis is correct, such design can efficiently reduce the asymmetry of information caused to re-matching.<sup>11</sup>

### 3.2.2 Experimental procedures.

The experiment involves a total of 12 sessions (6 for each experimental condition), each comprising 8 subjects. I use round-robin matching protocol (that assures that every two subjects have an opportunity to interact during the experiment) and a random group survival rule outlined above. Consequently, the structure of group matching and the number of rounds differ between sessions.

---

<sup>8</sup>Instructions (translated from French to English) are provided as a supplementary material.

<sup>9</sup>One may argue that such design also gives rise to (uncontrollable) effects caused by reputation-building. However, even if subjects indeed respond to an incentive to build a reputation, this incentive remains constant in the entire experiment: across different kinds of pairs – newly established and preserved, as well as under different information conditions. Yet, this symmetricity is another important refinement with respect to the usual partner-stranger comparison where reputation-building behavior may unobservedly arise in the former scheme, but not in the latter.

<sup>10</sup>This closely resembles the procedure adopted by Masclet, Noussair, Tucker, and Villeval (2003).

<sup>11</sup>At the same time, one may think that it also creates a potential risk of contamination of subject's behavior can be affected by the knowledge about other *pairs*. However, such risk does not seem substantial given that players *never* learn about the *entire* context of events occurring outside their own groups – that is, all members' contributions and corresponding disapproval points, which should preclude the possibility of inference about the patterns of behavior and evaluation in other groups.

In order to control for the effects of these variations, I use six independent, randomly generated matching sequences (henceforth labeled *Game 1*, ..., *Game 6*) and run two separate sessions for each of them: one implementing BC and one implementing EC.<sup>12</sup> Subjects are informed that the game contains between 10 and 16 rounds and that its length is determined randomly. In practice, sessions contain between 11 and 15 rounds, and pairs of subjects interact for up to five consecutive rounds.

At the beginning of each session, participants are randomly assigned to their computers and asked to fill in a small personal questionnaire containing basic questions about their age, gender, education, *etc.* The written instructions are then read aloud. Before starting, subjects are also asked to fill in a quiz assessing their understanding of the game they are about to play. Once the quiz and all remaining questions are answered, the experiment begins.

Once all pairs complete a round of the game, subjects are either informed that a new round starts, or that the experiment ends. In the latter case, a single round is randomly drawn and each participant receives the amount in EUR corresponding to his gains in that round (converted from ECU to EUR using an exchange rate  $1 \text{ ECU} = 0.40 \text{ EUR}$ ), plus a show-up fee equal to 5 EUR.

All sessions took place in the lab of University Paris 1 (LEEP) in July 2012. The recruitment of subjects has been carried out via LEEP database among individuals who have successfully completed the registration process on Laboratory's website.<sup>13</sup> Among 96 participants, 51 are males and 45 are females. 63 participant are students, among which 67% might have some background in game theory due to their field of study.<sup>14</sup> 82 subjects took part in experiments organized in LEEP in the past. Participants' average age is roughly 25. No subject participated in more than one experimental session. Each session lasted about 45 minutes, with an average payoff of 12.20 Euros.

### 3.3 Results

#### 3.3.1 Behavior with and without ex post communication

Table 3.1 presents subjects' average contributions as a function of the structure of experimental game and experimental condition. In five experimental games out of six, the presence of *ex post* communication increases the average contribution. This shift of behavior is significant at the 5% level according to the Wilcoxon signed-rank test.<sup>15</sup> <sup>16</sup> On the other hand, a Wilcoxon-Mann-

---

<sup>12</sup>The exact form of these sequences can be found in the supplementary material.

<sup>13</sup>The recruitment uses ORSEE Greiner (2004); the experiment is computerized through a software developed under REGATE Zeiliger (2000).

<sup>14</sup>Disciplines such as economics, engineering, management, political science, psychology, mathematics applied in social science, mathematics, computer science, sociology, biology.

<sup>15</sup>For each of the six games, this test matches average contributions observed in BC and in EC, and therefore accounts for the effects related to different game structures.

<sup>16</sup>All *p*-values used in statistical analysis correspond to two-sided tests.

Table 3.1: Average contributions according to treatment and experimental game

Conditions	Game 1	Game 2	Game 3	Game 4	Game 5	Game 6	Average	$p$
Average contributions in round 1								
Baseline	7.625	6.750	8.125	6.375	5.750	3.875	6.417	0.210
Evaluation	6.250	8.250	6.500	8.375	7.875	8.750	7.667	
Overall average contributions								
Baseline	5.000	2.900	4.420	3.083	6.010	2.846	3.942	0.046
Evaluation	4.083	7.783	6.045	6.325	9.135	8.029	6.938	
Session details								
Number of subjects	8	8	8	8	8	8		
Number of rounds	12	15	11	15	12	13		

**Note.** Columns 1-6 present average contributions in each experimental game, using data from round 1 (upper part) and all rounds (middle part). The last two columns summarize these results and provide non-parametric tests for the significance of the effect of treatment on contributions: the Wilcoxon-Mann-Whitney rank-sum test using individual observations in round 1, and Wilcoxon signed-rank test using game-level matched averages for the aggregate data. The lower part of the table contains additional information on the number of subjects and the length of each experimental game.

Whitney rank-sum test using individual observations from round 1 does not reject the hypothesis that subjects' behavior is the same in both experimental environments ( $p=0.210$ ). This allows us to formulate our first result:

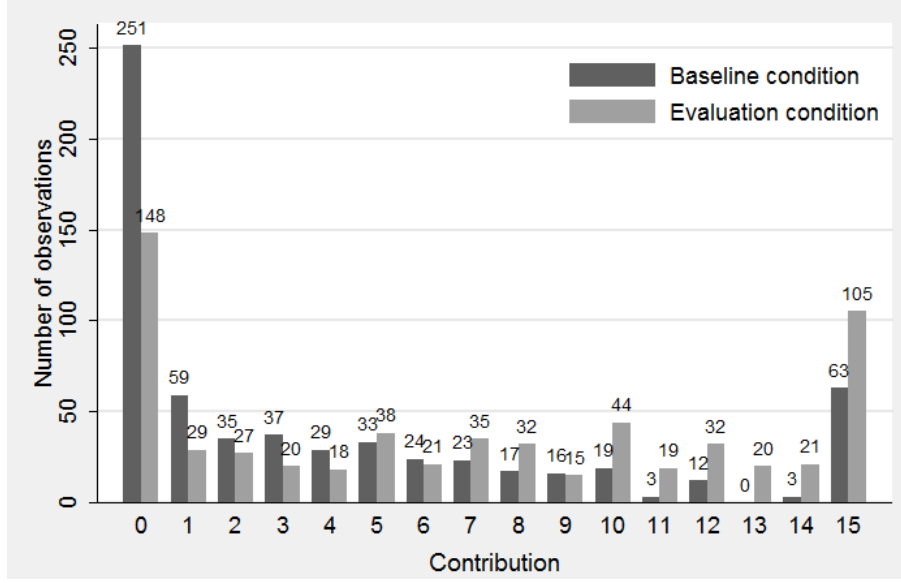
**Result 1.** The average contribution in the evaluation condition significantly increases as compared to the baseline condition. However, this effect does not occur in the initial round.

To further highlight the behavioral transition caused by the presence of communication, Figure 3.1 compares distributions of contributed amounts in both experimental conditions. In BC, over 40% (251 out of 624) of decisions result in a null contribution, as compared to less than 25% (148/624) in EC. Moreover, contributions between 1 ECU and 4 ECU are more frequent in the former than in the latter. Then, the relationship between the two conditions is unstable until the threshold level of 10 ECU, above which all values appear substantially more often in EC than in BC, including the case in which the entire endowment is transferred to the common pool (105 times in EC against 63 times in BT). In addition, Figure 3.2 suggests that this phenomenon is robust to the way pairs are matched prior to interacting. Within each experimental condition, one observes similar distributions of contributions regardless of whether pairs have been maintained or changed before interacting. On the other hand, the key difference between the condition evoked in Figure 3.1 prevails: under *ex post* communication the frequency of low contributions is substantially lower than without communication, while the opposite holds for high contributions.<sup>17</sup>

To formalize this insight, I estimate parametric regression models that provide further statis-

<sup>17</sup>For instance, null contributions are observed 134 in BC in the absence of prior re-matching of pairs, and 110 should re-matching occur. Analogous figures in EC are 80 and 65. Moving to the opposite extreme, contributions of 15 occur 31, 26, 54 and 42 times in the four respective cases.

Figure 3.1: Distribution of contributions across treatments

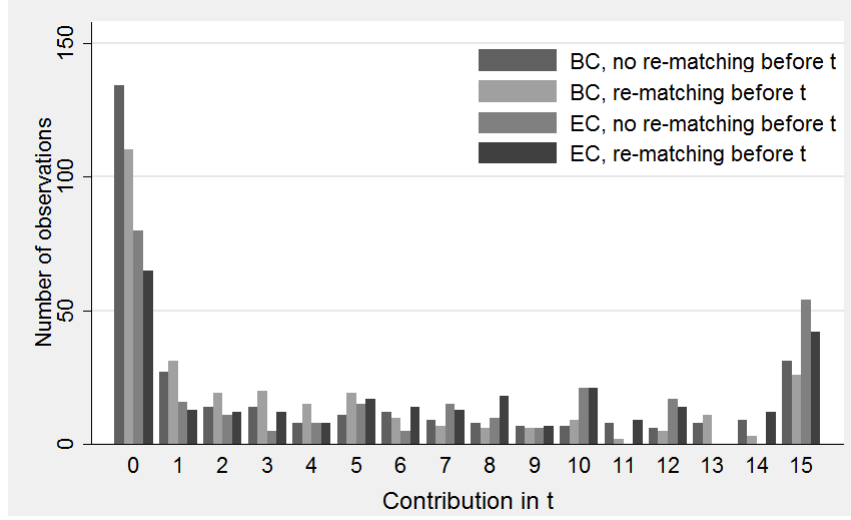


**Note.** For each experimental treatment, data contain 624 observations from 6 experimental sessions. Contributions are given in ECU.

tical evidence for the main forces driving the behavioral transition depicted in Figures 3.1 and 3.2. The core instruments for this analysis are embedded in the  $2 \times 2$  design of this experiment: the first dimension being the presence of random  $1_{Re-matching_{t-1}^t}$  (set to 1 if re-matching occurs, 0 otherwise) of pairs prior to round  $t$ , while the second – the presence of *ex post* communication treatment ( $1_{EC} = 1$  for evaluation condition, 0 otherwise). The statistical analysis focuses on two distinct outcomes: player  $i$ 's own contribution at time  $t$  ( $c_t^i$ ) and the absolute difference between the contribution of player  $i$  and player  $j$  who form a pair at time  $t$  ( $|c_t^i - c_t^j|$ ).

The intuition underlying the utilization of these variables is the following. As far as the first outcome is concerned, the signaling role of *ex post* communication may be confirmed by the existence of a significant impact of  $1_{EC}$  on contributions in maintained pairs, and a significant difference between contributions observed in maintained and re-matched pairs when  $1_{EC} = 1$  – both of which jointly indicate that the behavioral effect of *ex post* communication draws on the continuity of interaction. In addition, a significant effect of  $1_{EC}$  among newly created pairs suggests that the presence of *ex post* communication also affects subjects' choices when interactions are not continuous, which points to the importance of the emotion-based hypothesis. This explanation becomes a dominant one in the absence of a significant difference between contributions in maintained and re-matched pairs when  $1_{EC} = 1$ . These mechanisms are tested in Model 1. Importantly, the distributions presented in Figures 3.1 and 3.2 suggest that an important part of

Figure 3.2: Contributions, ex-post communication and re-matching



**Note.** For each combination of experimental condition (BC or EC) and matching structure (re-matching occurs before period  $t$  or not), data contain 288 observations from 6 experimental sessions. Contributions are given in ECU.

variations in behavior captured by Model 1 may result from variations in the frequency of extreme contributions, 0 and 15. To address this issue, Model 2 provides a robustness check by withdrawing these extreme observations from the sample. Model 3 uses dependent variable  $|c_t^i - c_t^j|$  to provide a complementary test for the signaling hypothesis. If *ex post* communication operates as an efficient signaling mechanism, then it should allow players to mutually adapt their strategies and attain more equilibrated outcomes within pairs, so that the outcomes observed in maintained groups playing under evaluation condition should involve less intra-group inequality than elsewhere.

The estimates of coefficients from these three models are summarized in Table 3.2.<sup>18</sup> In addition to dummy variables  $1_{Re-matching_{t-1}^t}$  and  $1_{EC}$  as well as their interaction, these models also control for time-related effects (dummy variables representing *Round 2*, ..., *Round 15*) as well as effects related to particular matching sequences (dummy variables representing *Game 1*, ..., *Game 6*). Model 1 delivers an important set of evidence favoring the emotional-based explanation for the effect of *ex post* communication in a repeated public goods game, while no statistical support for the strategic signaling hypothesis. First, the presence of communication has a positive impact on contributions in both maintained ( $H_0 : \beta_2 = 0, p = 0.064$ ) and re-matched pairs ( $H_0 : \beta_2 + \beta_3 = 0, p = 0.028$ ). Second, players' behavior in the evaluation condition does not depend on whether groups were previously re-matched or not ( $H_0 : \beta_1 + \beta_3 = 0, p = 0.516$ ). Model 2 confirms the

<sup>18</sup>Marginal effects corresponding to these regressions are provided as a supplementary material, and provide further statistical support for treatment effects discussed in this section.

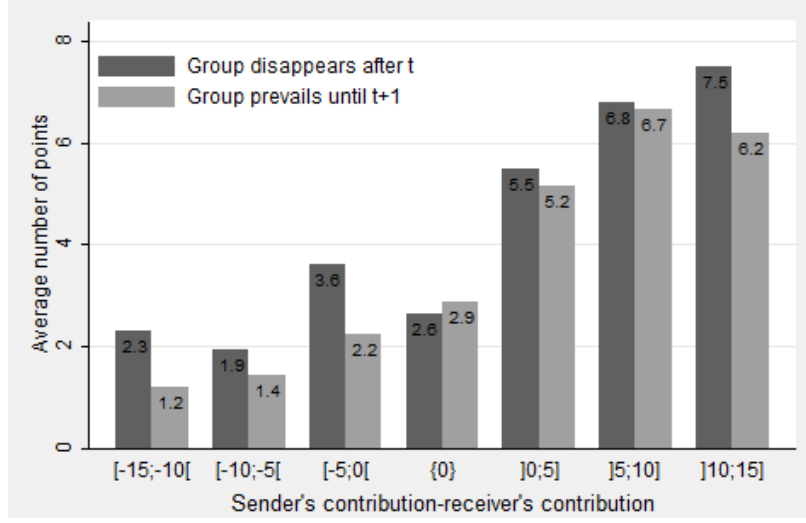
Table 3.2: Determinants of individual and group behavior

Dep. variable:	Model 1 $c_t^i   c_t^i \in [0, 15]$		Model 2 $c_t^i   c_t^i \in [1, 14]$		Model 3 $ c_t^i - c_t^j $	
	$\beta$	$p_\beta$	$\beta$	$p_\beta$	$\beta$	$p_\beta$
<i>Intercept</i> ( $\beta_0$ )	1.406	0.815	3.131	0.019	3.534	0.221
$1_{Re-matching_{t-1}^t}$ ( $\beta_1$ )	1.202	0.119	-0.240	0.619	1.659	0.072
$1_{EC}$ ( $\beta_2$ )	5.398	0.064	3.047	0.046	1.229	0.386
$1_{EC} \times 1_{Re-matching_{t-1}^t}$ ( $\beta_3$ )	-0.600	0.597	0.403	0.480	-1.311	0.351
Additional controls:						
<i>Round</i> :						
3	-1.683	0.072	0.230	0.840	1.979	0.228
4	-0.720	0.679	-0.515	0.564	0.911	0.555
5	-2.233	0.058	-0.969	0.147	0.537	0.609
6	-1.880	0.042	-0.695	0.479	0.708	0.574
7	-2.903	0.010	-0.241	0.813	0.061	0.960
8	-3.593	0.024	-0.761	0.512	-1.359	0.308
9	-3.406	0.008	-0.770	0.335	0.469	0.782
10	-4.491	0.015	-1.642	0.129	-1.614	0.360
11	-4.083	0.023	-1.190	0.316	0.434	0.770
12	-4.059	0.006	-1.674	0.117	0.602	0.714
13	-4.367	0.035	-1.061	0.360	0.508	0.764
14	-3.648	0.044	-0.605	0.641	1.326	0.522
15	-4.772	0.031	0.015	0.988	0.501	0.629
<i>Game</i> :						
2	1.347	0.847	2.923	0.012	-0.688	0.790
3	1.718	0.771	0.579	0.846	-1.603	0.557
4	1.409	0.811	1.118	0.256	-1.139	0.699
5	5.827	0.325	3.597	0.012	-0.328	0.899
6	1.994	0.750	2.549	0.342	-1.265	0.630
N	1152		610		576	

**Note.** Estimates come from double-censored tobit regressions. Dependent variable in Model 1 is player  $i$ 's contribution at time  $t$  ( $c_t^i$ ), the same variable conditional on not taking an extreme value (0 or 15) in Model 2, whereas Model 3 uses the absolute difference between the contributions of player  $i$  and player  $j$  forming a group at time  $t$  ( $|c_t^i - c_t^j|$ ). Explanatory variables include a dummy indicating the occurrence of group re-matching prior to ( $1_{Re-matching_{t-1}^t}$ ) and another dummy indicating whether the game involves *ex post* communication ( $1_{EC}$ ), as well as their interactions. Models also control for time and matching-sequence effects. Model 1 (2) uses individual-level (group-level) observations. Data cover observation from round 2 onwards. Variables *Round 2* and *Game 1* serve as reference points and are omitted in regressions. Standard errors are clustered at the session level (12 clusters) and are corrected using the leave-one-out jackknife procedure.

robustness of these findings once the extreme observations (*i.e.* contributions equal to 0 or 15)

Figure 3.3: Sent disapproval points and the relative size of contribution



**Note.** Data come from 6 sessions and contain 288 observations for each type of group. Contributions are given in ECU.

are eliminated from the sample.<sup>19</sup> Model 3 provides evidence in the same vein. One cannot reject the joint hypothesis that the level of inequality in maintained pairs is neutral to the *ex post* communication variable, and that the level of inequality in the evaluation condition depends on the nature of groups' re-matching ( $H_0 : \beta_2 = 0 \cap \beta_1 + \beta_3 = 0, p = 0.272$ ). Interestingly, in the baseline condition the contributions within newly formed group are more divergent than in retained group ( $H_0 : \beta_2 = 0, p = 0.072$ ), which suggests that the experience acquired in the course of a continuous interaction improves players' capacity of strategic adaptation, and that this effect is superseded by the *ex post* exchange of information. This yields the second result established in this section:

**Result 2.** Allowing for exchange of *ex post* messages restrains opportunistic behavior in both kinds of groups: newly formed pairs and pairs that continue interactions from the past. Furthermore, in both experimental conditions average contributions do not vary significantly between newly established and retained groups. Finally, the intra-group inequality of contributions in the evaluation condition is not significantly greater in newly established pairs than in preserved ones, unlike the analogous inequality in the baseline condition.



### 3.3.2 Formation of ex-post messages

Results reported so far show that the institution of *ex post* communication contributes to reducing the scope of opportunistic behavior. This section addresses two central questions related to the formation of *ex post* messages: (i) what are the patterns of relationship between *ex post* messages and underlying contributions, and (ii) do these patterns vary between retained and re-matched groups? The answer to the first question is provided by Figure 3.3 which presents the average number of points sent by subject to his group member as a function of his deviation from other group member's contribution, and the outcome of the random draw that determines group's future. Under both matching scenarios, the overall pattern is coherent: the number of sent disapproval points grows as the difference between sender's and receiver's contributions grows.

Answering the second question requires a more formal insight into the process of message formation. For this sake, I estimate a linear model representing the number of points received by a player in round  $t$  as a function of his *Group\_member's\_contribution<sub>t</sub>*, as well as its relative size within his group – i.e. the magnitude of *Sender's\_positive\_deviation<sub>t</sub>* or *Sender's\_absolute\_negative\_deviation<sub>t</sub>* with respect to other group member's decision. The effect of prospective re-matching is captured by the variable  $1_{Re-matching_t^{t+1}}$  (like before, set to 1 if re-matching is about to occur between  $t$  and  $t + 1$ , a 0 otherwise) as well as its interactions with the above variables.

In the absence of re-matching prior to the upcoming interaction, players award more disapproval points the less their group members contribute ( $H_0 : \gamma_1 = 0, p = 0.003$ ), and the more they contribute themselves relative to their group members ( $H_0 : \gamma_2 = 0, p = 0.008$ ). The magnitude of own *negative* deviation (in absolute terms), in turn, does not play a significant role in the process of point attribution ( $H_0 : \gamma_3 = 0, p = 0.463$ ). Once re-matching occurs, the significance of these variables does not change with an exception of *Sender's\_positive\_deviation<sub>t</sub>* which becomes slightly insignificant ( $H_0 : \gamma_1 + \gamma_4 = 0, p = 0.005, H_0 : \gamma_2 + \gamma_5 = 0, p = 0.119, H_0 : \gamma_3 + \gamma_6 = 0, p = 0.365$ , respectively). Importantly, one may not reject the joint hypothesis that the prospect of re-matching does not affect the pattern by which group members' contributions affect transmitted disapproval points ( $H_0 : \gamma_4 = 0 \cap \gamma_5 = 0 \cap \gamma_6 = 0, p = 0.471$ ). This leads us to the following result:

**Result 3.** *Ex post* disapproval points are attributed in an informative and coherent manner: low contributors receive more disapproval points than high contributors. This pattern of point attribution does not vary significantly between pairs that cease and continue interacting.

### 3.4 Summary and discussion

An important body of economic literature analyses the impact of communication and emotions on behavior in economic problems involving conflicting interests. For instance, Ellingsen and Johannesson (2004) and Vanberg (2008) argue that *ex ante* communication reduces the scope of

---

<sup>19</sup>The  $p$ -values corresponding to the three respective tests are: 0.046, 0.025, 0.721.

Table 3.3: Sent disapproval points and the relative size of contribution

Determinants:	$\gamma$	$p$
<i>Intercept</i> ( $\gamma_0$ )	7.831	0.001
<i>Group_member's_contribution<sub>t</sub></i> ( $\gamma_1$ )	-0.766	0.003
<i>Sender's_positive_deviation<sub>t</sub></i> ( $\gamma_2$ )	0.314	0.008
<i>Sender's_absolute_negative_deviation<sub>t</sub></i> ( $\gamma_3$ )	0.128	0.463
$1_{Re-matching_t^{t+1}}$ ( $\gamma_4$ )	-1.113	0.400
$1_{Re-matching_t^{t+1}} \times \textit{Group\_members's\_contribution}_t$ ( $\gamma_5$ )	0.066	0.713
$1_{Re-matching_t^{t+1}} \times \textit{Sender's\_positive\_deviation}_t$ ( $\gamma_6$ )	0.048	0.752
$1_{Re-matching_t^{t+1}} \times \textit{Sender's\_absolute\_negative\_deviation}_t$ ( $\gamma_7$ )	-0.220	0.249
Number of obs.	576	

**Note.** Double-censored tobit regressions of the number of points sent in  $t$  on variables indicating the level of *Group\_member's\_contribution<sub>t</sub>*, the value of *Sender's\_positive\_deviation<sub>t</sub>* or an *Sender's\_absolute\_negative\_deviation<sub>t</sub>*, a dummy indicating the occurrence of  $1_{Re-matching_t}$  as well as its interactions with the three above variables. Standard errors are estimated using observations clustered at the session level (6 clusters in total), and then corrected using delete-one jackknife.

opportunistic behavior due to agents' *aversion to lying*, while Charness and Dufwenberg (2006) relate this transition to *guilt aversion* due to which agents experience disutility from letting down others' expectations. Echoing the previous findings by Masclet, Noussair, Tucker, and Villeval (2003) and Dugar (2013), this paper reports that the availability of *ex post* communication involving costless evaluation points substantially reduces opportunistic behavior in a repeated VCM game. More importantly, it offers a new angle for understanding this important behavioral process. The experimental design aims at highlighting two potential (and non-excludable) factors behind the behavioral effect of *ex post* communication: first, the preference for approval/aversion to disapproval, understood as an emotional utility or disutility drawn from others' opinions about own behavior; second, the transmission of strategic signals (such as threats) linked to future interactions. Existing experimental literature lacks a clear-cut consensus on the role and the extent of these effects. Masclet, Noussair, Tucker, and Villeval (2003) suggest that both of these mechanisms may influence subjects' behavior, Peeters and Vorsatz (forthcoming) explain their results in terms of strategic information transmission, while the interpretation of Dugar (2013) points towards the importance of emotions induced by others' judgments about own actions.

Data from this experiment tend to support the emotion-based explanation. First, I observe similar behavioral effects of *ex post* communication in both kinds of groups: those who are about to cease interacting (and thus may not exchange strategic signals about their future decisions) and those who continue interacting over time (and thus may exchange future-oriented strategic information). Second, the relationship between contributions and the resulting level of disapproval

is coherent (low contributors face stronger disapproval than high contributors) and does not vary as a function of groups' fate. Notwithstanding the strategic signaling explanation, *ex post* communication does not improve players' strategic adaptation (measured as the within-group inequality in contributions), even in interactions that extend throughout multiple periods.

In a recent study, Charness and Dufwenberg (2010) discuss two implementations of pre-play, cheap-talk, natural language communication in a trust game: an unstructured (free-form) protocol where (almost) any content may be transmitted between players, and a fixed-form protocol which only allows for single-phrase, standardized "bare" promises. Their conclusion is that only the unrestricted communication procedure induces emotional reaction from subjects (which, in their study, is explained as a personal cost caused by lying-aversion). The main result pinned down by this study differs at face value from Charness and Dufwenberg's insight. In fact, even an structured, artificial and wordless method of expressing own judgment – such as the utilization of evaluation points – creates an environment in which the aversion to others' disapproval refrains opportunistic behavior in a social dilemma. On the other hand, a closer comparison of experimental investigations using such communication protocol and experiments utilizing communication based on natural language points to some important differences between these two mechanisms. The present study reports no significant treatment effect on average contributions in the initial round, and important behavioral variations in subsequent periods. This may suggest that the preference for *avoiding* disapproval is related to the previous *exposure* to feedback, so that the effectiveness of this structured mechanism is not predetermined, but rather involves learning of conventions and arises via procedural experience. Moreover, this phenomenon persists in previous lab experiments using this communication procedure (Dugar, 2010, 2013). Different conclusions are drawn from studies using natural language communication (of unrestricted or fixed form), in which the mere *anticipation* of others' feedback about own actions already suffices to make subjects behave so as to *avoid* disapproval (Ellingsen and Johannesson, 2008; Xiao and Houser, 2009; López-Pérez and Vorsatz, 2010). This issue seems particularly relevant in certain real life implementations, such as the selection of the feedback scheme for an e-commerce platform. A more systematic comparison of different feedback mechanisms would certainly constitute a desirable avenue for future research.

## Supplementary material

### Written instructions

*Author's note: BC (EC) at the beginning of a paragraph indicates that this part is specific to baseline condition (evaluation condition)*

You are about to take part in an experiment in which you can earn money. Your gains will depend on your decisions, as well as on the decisions made by other participants.

Before starting, we would like to ask you to answer a few standard questions (concerning your age, education, profession, ...) that will help us to get to know you better. **This information, as well as the amount of your gains in this experiment, will remain strictly confidential and anonymous.**

Please, fill in the questionnaire using the interface on your computer screen, which is divided into three parts:

- In the *top* section, you will find information that might help you in making decisions.
- In the *middle* section, you will submit your decisions by clicking on a relevant button.
- In the *bottom* section, you will see all your decisions and gains from previous rounds of the experiment.

**Thank you.**

## THE EXPERIMENT

The experiment consists of several rounds. The total number of rounds is random and might vary between 10 and 16. In each round, participants are divided into groups of two. In each round (more precisely, in Stage 2 described below) a random draw decides that:

- either your group will not change in the next round;
- or that your group will change in the next round. In order to form your new group, another participant will be chosen at random among participant who have never been part of your group before.

**Both events are equally probable – each occurs with a 50% chance.**

In every round, each participant's gain is determined as follows. At the beginning of round, every person possesses the initial endowment of 15 ECU (Experimental Currency Unit).

Members of each group may create a common pool. Each participant freely chooses his level of contribution to the pool that may range between 0 ECU and 15 ECU. The total amount gathered in the pool is then multiplied by 1.5 and divided equally among group members.

For instance, if you are participant  $i$  who contributed  $N_i$ , and the other member of your group, participant  $j$ , contributed  $M_j$ , then your gain ( $Gain_i$ ) equals:

$$Gain_i = 15 - 0.25 \times N_i + 0.75 \times M_j \quad (3.2)$$

The Table provided below presents your gain in ECU in a given period as a function of your level of contribution and the other group member's level of contribution.

Table 3.4: Your gain in ECU in a given period as a function of your decision and your group member's decision

		Your group member's level of contribution															
		0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Your level of contribution	0	15	15.75	16.5	17.25	18	18.75	19.5	20.25	21	21.75	22.5	23.25	24	24.75	25.5	26.25
	1	14.75	15.5	16.25	17	17.75	18.5	19.25	20	20.75	21.5	22.25	23	23.75	24.5	25.25	26
	2	14.5	15.25	16	16.75	17.5	18.25	19	19.75	20.5	21.25	22	22.75	23.5	24.25	25	25.75
	3	14.25	15	15.75	16.5	17.25	18	18.75	19.5	20.25	21	21.75	22.5	23.25	24	24.75	25.5
	4	14	14.75	15.5	16.25	17	17.75	18.5	19.25	20	20.75	21.5	22.25	23	23.75	24.5	25.25
	5	13.75	14.5	15.25	16	16.75	17.5	18.25	19	19.75	20.5	21.25	22	22.75	23.5	24.25	25
	6	13.5	14.25	15	15.75	16.5	17.25	18	18.75	19.5	20.25	21	21.75	22.5	23.25	24	24.75
	7	13.25	14	14.75	15.5	16.25	17	17.75	18.5	19.25	20	20.75	21.5	22.25	23	23.75	24.5
	8	13	13.75	14.5	15.25	16	16.75	17.5	18.25	19	19.75	20.5	21.25	22	22.75	23.5	24.25
	9	12.75	13.5	14.25	15	15.75	16.5	17.25	18	18.75	19.5	20.25	21	21.75	22.5	23.25	24
	10	12.5	13.25	14	14.75	15.5	16.25	17	17.75	18.5	19.25	20	20.75	21.5	22.25	23	23.75
	11	12.25	13	13.75	14.5	15.25	16	16.75	17.5	18.25	19	19.75	20.5	21.25	22	22.75	23.5
	12	12	12.75	13.5	14.25	15	15.75	16.5	17.25	18	18.75	19.5	20.25	21	21.75	22.5	23.25
	13	11.75	12.5	13.25	14	14.75	15.5	16.25	17	17.75	18.5	19.25	20	20.75	21.5	22.25	23
	14	11.5	12.25	13	13.75	14.5	15.25	16	16.75	17.5	18.25	19	19.75	20.5	21.25	22	22.75
	15	11.25	12	12.75	13.5	14.25	15	15.75	16.5	17.25	18	18.75	19.5	20.25	21	21.75	22.5

## WHAT HAPPENS IN EACH ROUND

At the beginning of each round (accept for the first one), the result of the random draw that took place in **the previous period** is recalled to each participant. Each participant is informed that:

- either his group remains the same as in the previous round;
- *BC*: or in the ongoing round he will play with a different person. In this case, each participant is also informed about the level of contribution of his current group member in the previous period.
- *EC*: or in the ongoing round he will play with a different person. In this case, each participant is also informed about the level of contribution of his current group member and the number of points he received in the previous period.

Every round contains 3 stages:

**Stage 1.** Each participant chooses his level of contribution to the pool.

**Stage 2.** Each participant observes the outcomes of a random draw that determines his group in the next round.

**Stage 3.** *BC*: Finally, every participant is informed about his group member's level of contribution and his own gain for the round.

*EC*:

- At the beginning of this stage, every participant is informed about his group member's level of contribution and his own gain for the round.
- Then, each participant has to possibility to express his opinion about the other group member by assigning him a certain number of points. **A high number of points expresses disapproval: 10 points correspond to the strongest disapproval, 0 points correspond to the weakest disapproval.** To do this, you should select the number of points on your screen and press OK button to confirm your choice. The choice of the number of points **has no impact on either participants' gains in the experiment.**
- Finally, **every participant is informed about the number that were assigned to him** by the other group member.

At the end of each round, a message on your screen will inform you that either a new round is about to start, or that the experiment ends.

## **PAYMENT OF YOUR EARNINGS**

At the end of the experiment, **one round is selected at random**. Each participant receives a sum in EUR corresponding to the amount he earned in this round, converted from ECU to EUR using an exchange rate  $1 \text{ ECU} = 0.40 \text{ EUR}$ , plus a bonus of 5 € for completing the experiment. All payments are made individually and in cash.

For obvious reasons, **you are not allowed to talk during the experiment**. Participants who violate this rule will be excluded from the experiment and all payments. It is crucial that you understand perfectly the rules of this experiment. Should you have any questions to ask, please raise your hand.

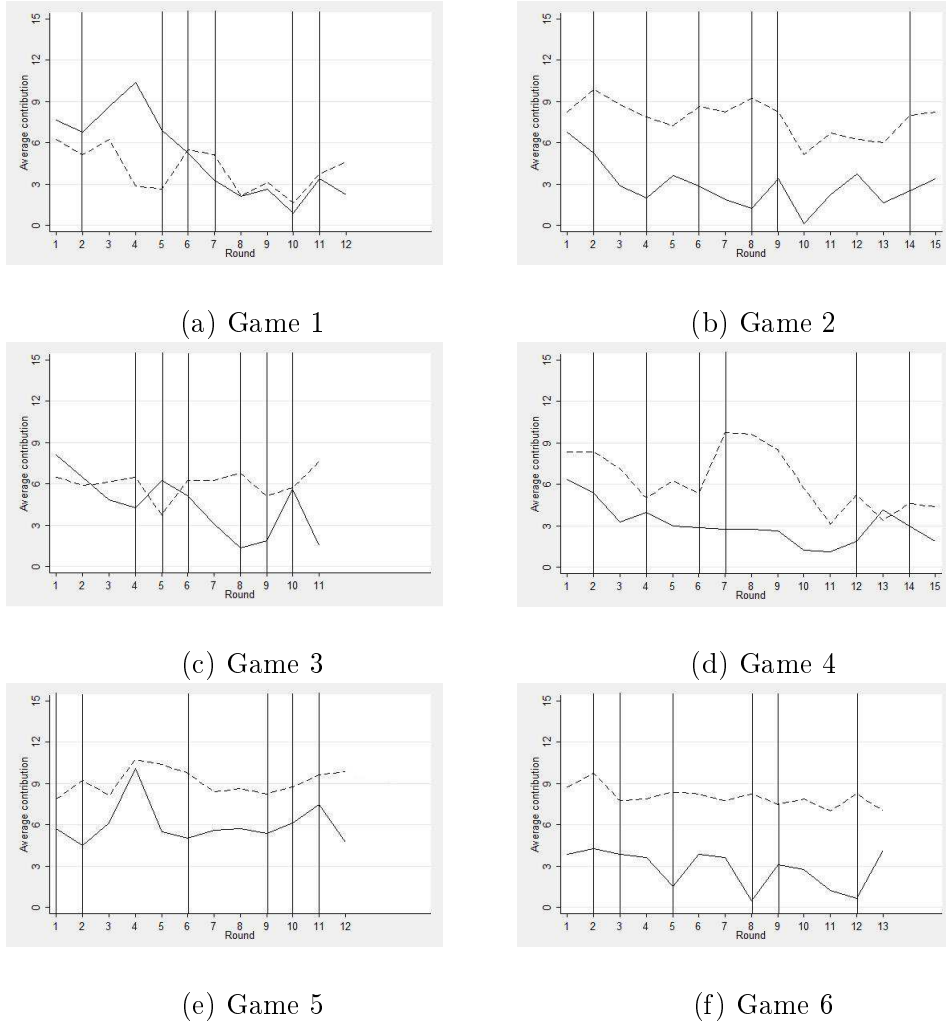
**Thank you for your participation!**



## Evolution of contributions in experimental games

Figures 3.3a-f suggest that treatment effect is stable over time: in Games 2, 5, and 6 contributions in evaluation condition dominate those in baseline condition in all rounds; in Games 3 and 4 this tendency is reversed in only 3 rounds out of 11, and 1 round out of 15, respectively. Solely in Game 1 the relationship is more ambiguous.

Figure 3.3: Evolution of average contribution by game and treatment



**Note.** Vertical lines indicate rounds after which groups are being re-matched. Solid line indicates baseline condition, dashed line indicates evaluation condition.

## Marginal effects for regression models in Table 4.2

Table 3.5 reports the (average) marginal effects of explanatory variables corresponding to regression models presented in Section 3.3.1 (Table 3.2). These results confirm the statistical significance of the treatment effects reported in that section.

Table 3.5: Determinants of individual and group behavior: marginal effects

Dep. variable:	Model 1 $c_t^i   c_t^i \in [0, 15]$		Model 2 $c_t^i   c_t^i \in [1, 14]$		Model 3 $ c_t^i - c_t^j $	
	ME	$p_{ME}$	ME	$p_{ME}$	ME	$p_{ME}$
$1_{Re-matching_{t-1}^t}$	0.646	0.083	-0.196	0.611	1.184	0.046
$1_{EC}$	2.901	0.041	2.480	0.019	0.877	0.375
$1_{EC} \times 1_{Re-matching_{t-1}^t}$	-0.323	0.588	0.328	0.470	-0.935	0.332
Additional controls:						
<i>Round :</i>						
3	-0.957	0.045	0.192	0.836	1.467	0.186
4	-0.413	0.669	-0.426	0.552	0.659	0.541
5	-1.263	0.027	-0.796	0.125	0.385	0.592
6	-1.067	0.018	-0.573	0.463	0.510	0.556
7	-1.628	0.001	-0.200	0.808	0.043	0.96
8	-1.996	0.004	-0.627	0.501	-0.916	0.292
9	-1.897	0.001	-0.635	0.318	0.335	0.776
10	-2.458	0.001	-1.332	0.112	-1.077	0.333
11	-2.251	0.006	-0.974	0.294	0.310	0.764
12	-2.238	0.001	-1.357	0.087	0.433	0.708
13	-2.396	0.009	-0.870	0.332	0.364	0.76
14	-2.025	0.018	-0.500	0.627	0.969	0.509
15	-2.599	0.01	0.0122	0.987	0.359	0.615
<i>Game :</i>						
2	0.695	0.842	2.361	0.001	-0.504	0.787
3	0.894	0.757	0.439	0.845	-1.146	0.554
4	0.728	0.8	0.863	0.21	-0.825	0.695
5	3.227	0.26	2.934	0.002	-0.242	0.898
6	1.044	0.736	2.046	0.336	-0.913	0.628
N	1152		610		576	

# General conclusion

Economic literature often perceives communication in collective decision-making problems as a simple way of transmitting strategic information between rational economic agents, who then use it to maximize their utility. From this perspective, the effect of communication depends primarily on all parties' strategic goals.

Experimental evidence presented in this thesis suggests that the nature of communication is not limited to strategic information transmission. In particular, it highlights the importance of two psychological factors underlying the effect of communication on human subjects' behavior: personal commitment and aversion to others' expressed disapproval.

## Communication and commitment

In Chapter 1, we discuss a multilateral decision-making problem in which it is possible to simultaneously maximize all parties' monetary gains. However, numerous lab implementations show that players frequently fail to benefit from this compatibility of strategic interests. Suboptimal outcomes arise due to two different reasons: first, players are very frequently reluctant to rely on others' preference to maximize own payoff, especially when their potential surplus from doing so is not substantial; second, the payoff-maximizing behavior indeed happens not to be an absolutely dependable one.

Our investigation aims at dealing with this inefficiency by allowing subjects to exchange strategic signals before taking actions. While both communication and an alternative source of information – observation of other players' past behavior – are found to facilitate strategic adaptation by reducing players' uncertainty about others' behavior, neither of them suffice to render this behavior more efficient. As a result, both mechanisms involving transmission of information help overcome coordination failure, but fall short of improving the overall efficiency of outcomes. In a broader perspective, this experiment identifies important limitations of cheap-talk communication – a mechanism generally considered as a useful means to improve the quality of economic interactions (Crawford, 1998). In the absence of a pronounced link between one's words and actions, institutions involving communication may well happen to be insufficient for this purpose.

In Chapter 2, we address this issue using the social psychological theory of commitment.

We aim at enhancing commitment underlying cheap-talk communication by engaging agents into a truth-telling behavior by means of a voluntary oath to tell the truth. Results are stunning: by enhancing the trustworthiness of and the trust in transmitted messages, fostering efficient behavior and eliminating strategic uncertainty, oath induces a substantial (50%) improvement in the efficiency of outcomes.

Altogether, this work demonstrates that a simple non-monetary pre-commitment device – communication under truth-telling oath – may provide a powerful means to improve the efficiency of human interactions.

## Communication and aversion to disapproval

Chapter 3 explores the sources of a positive behavioral effect of *ex post* communication in repeated interactions involving social dilemmas. Experimental design offers a way of separating two potential (and non-excludable) channels by which this institution may affect decisions: strategic information transmission and emotional sanctioning. Collected data speak in favor of the latter mechanism, suggesting that pro-social behavior is fostered by subjects' aversion to being disapproved by others. Importantly, this experiment also suggests that inducing an emotional response in humans does not necessarily require communicating in natural language, but may well occur in an environment with artificial and wordless signals, such as ordered numbers expressing the degree of disapproval. This stands at odds with the conclusions from some of the previous experiments (Charness and Dufwenberg, 2010).

## Open avenues for future research

Results provided in the present thesis open at least two interesting directions for future research.

First, existing economic literature puts forward mechanisms of pre-commitment that rely on a typically "economic" notion of pecuniary engagement. Bracht and Feltovich (2008) study a trust game with a monetary pre-commitment procedure: an escrow to be forfeited as a consequence of opportunistic behavior.<sup>20</sup> Escrow is found to significantly affect behavior and outcomes. In particular, the efficiency of interaction (which requires trust from one player and trustworthiness from the other) is high in the environment where monetary consequences of opportunistic behavior are steep.

---

<sup>20</sup>Trust game is a simple collective-action game played between two players: the trustor has the choice of investing or not investing in a project, which is administered by the trustee. The investment is successful, in the sense that the amount invested multiplies in value. However, the trustee controls the proceeds of investment: he may keep the total amount for himself or split it evenly with the trustor. The prediction of game theory for this game is dismal indeed: the unique subgame perfect equilibrium has the trustor refusing to invest, foreseeing that the trustee would keep the proceeds of investment. This equilibrium is inefficient, since total payoffs are higher if the trustor invests, and it is possible for the trustee to split the proceeds so that both players are strictly better off than in equilibrium.

Firm experimental results established in this thesis call for a thorough investigation of the role of pre-commitment in economic interactions. As an attempt to bridge the social psychological approach to pre-commitment that emphasizes personal engagement, and a more orthodox economic perspective based on monetary engagement, the next step of this investigation should involve a systematic "compare and contrast" experimental analysis of monetary and non-monetary pre-commitment mechanisms in collective-action economic problems.

Second, results included in this thesis raise the question of the dimensions of communication that cause emotional response in humans. Previous research suggests that the form of communication is of primary importance with this respect, and that two features foster emotional reaction: the affluence of content and the utilization of natural language. Evidence presented here differs at face value from these findings, pointing to an important emotional impact of an artificial and wordless communication mechanism. Unfortunately, economic literature still lacks a systematic investigation of this important dimension of communication in strategic interactions, a gap hopefully to be filled by the future research.

In a more general perspective, exploring the psychological factors underlying human communication in economic contexts may also result in interesting real life applications. For instance, recent years have witnessed an increasing popularity of websites on which one may engage to quit smoking, lose weight, exercise regularly, etc. This commitment is monitored by an external supervisor and it is also an option to back it with an online deposit of a certain sum of money. If the user fulfills his commitment, he gets back his bail, and loses it otherwise. One of the most popular websites of this kind, *Stickk.com*, reports that making a monetary engagement increases people's performance threefold. Thus, a better understanding of psychological mechanisms driving non-monetary commitment may yield interesting cost-efficient alternatives for users of such websites. Furthermore, exploring the relationship between emotions induced by communication and economic behavior may also help improve the design of feedback schemes for e-commerce.

# Bibliography

- ABADIE, A. (2002): “Bootstrap Tests for Distributional Treatment Effects in Instrumental Variable Model,” *Journal of the American Statistical Association*, 97(457), 284–292.
- ANDREONI, J., AND R. CROSON (2008): *Partners versus Strangers: Random Rematching in Public Goods Experiments* vol. 1 of *Handbook of Experimental Economics Results*, chap. 82, pp. 776–783. Elsevier.
- ANGRIST, J. D., AND V. LAVY (2009): “The Effect of High School Achievement Awards: Evidence from a Randomized Trial,” *American Economic Review*, 99(4), 1389–1414.
- ARROW, K. J. (1974): “Limited Knowledge and Economic Analysis,” *American Economic Review*, 64(1), 1–10.
- AUMANN, R. (1990): “Nash Equilibria Are Not Self-Enforcing,” in *Economic Decision-Making: Games, Econometrics, and Optimization: Contributions in Honor of Jacques H. Dreze*, ed. by J. J. Gabszewic, J.-F. Richard, and L. A. Wolsey, pp. 201–206. North-Holland, Amsterdam.
- BEARD, T. R., AND J. BEIL, RICHARD O. (1994): “Do People Rely on the Self-Interested Maximization of Others? An Experimental Test,” *Management Science*, 40(2), 252–262.
- BEARD, T. R., R. O. J. BEIL, AND Y. MATAGA (2001): “Reliant behavior in the United States and Japan,” *Economic Inquiry*, 39(2), 270–279.
- BELL, R. M., AND D. F. MCCAFFREY (2002): “Bias Reduction in Standard Errors for Linear Regression with Multi-Stage Samples,” *Survey Methodology*, 28(2), 169–179.
- BINMORE, K., J. MCCARTHY, G. PONTI, L. SAMUELSON, AND A. SHAKED (2002): “A Backward Induction Experiment,” *Journal of Economic Theory*, 104(1), 48–88.
- BLANCO, M., D. ENGELMANN, AND H. T. NORMANN (2011): “A within-subject analysis of other-regarding preferences,” *Games and Economic Behavior*, 72(2), 321–338.
- BLUME, A., AND A. ORTMANN (2007): “The effects of costless pre-play communication: Experimental evidence from games with Pareto-ranked equilibria,” *Journal of Economic Theory*, 132(1), 274–290.
- BOCHET, O., T. PAGE, AND L. PUTTERMAN (2006): “Communication and punishment in voluntary contribution experiments,” *Journal of Economic Behavior & Organization*, 60(1), 11–26.

- BOLTON, G. E., E. KATOK, AND A. OCKENFELS (2004): “How Effective Are Electronic Reputation Mechanisms? An Experimental Investigation,” *Management Science*, 50(11), 1587–1602.
- BOLTON, G. E., AND A. OCKENFELS (2006): “Inequality Aversion, Efficiency, and Maximin Preferences in Simple Distribution Experiments: Comment,” *American Economic Review*, 96(5), 1906–1911.
- BRACHT, J., AND N. FELTOVICH (2008): “Efficiency in the trust game: an experimental study of precommitment,” *Int. J. Game Theory*, 37(1), 39–72.
- BRACHT, J., AND N. FELTOVICH (2009): “Whatever you say, your reputation precedes you: Observation and cheap talk in the trust game,” *Journal of Public Economics*, 93(9-10), 1036–1044.
- BROCAS, I., J. CARRILLO, AND M. DEWATRIPONT (2003): “Commitment devices under self-control problems: An overview,” in *The Psychology of Economic Decisions*, ed. by I. Brocas, and J. Carillo, vol. 2, chap. 4, pp. 49–66. Oxford University Press, Oxford.
- BURGER, J. M. (1999): “The Foot-in-the-Door Compliance Procedure: A Multiple-Process Analysis and Review,” *Personality and Social Psychology Review*, 3(4), 303–325.
- CAMERER, C. F. (2003): *Behavioral Game Theory: Experiments in Strategic Interaction*. Princeton University Press, Princeton, NJ.
- CAMERON, A. C., J. B. GELBACH, AND D. L. MILLER (2008): “Bootstrap-Based Improvements for Inference with Clustered Errors,” *Review of Economics and Statistics*, 90(3), 414–427.
- CASON, T. N., AND D. FRIEDMAN (1997): “Price Formation in Single Call Markets,” *Econometrica*, 65(2), 311–345.
- ÇELEN, B., S. KARIV, AND A. SCHOTTER (2010): “An Experimental Test of Advice and Social Learning,” *Management Science*, 56(10), 1687–1701.
- CHARNESS, G. (2000): “Self-Serving Cheap Talk: A Test of Aumann’s Conjecture,” *Games and Economic Behavior*, 33(2), 177–194.
- CHARNESS, G., AND M. DUFWENBERG (2006): “Promises and Partnership,” *Econometrica*, 74(6), 1579–1601.
- (2008): “Broken Promises: An Experiment,” *UCSB Working Paper*, 10(08).
- (2010): “Bare promises: An experiment,” *Economics Letters*, 107(2), 281–283.
- CHARNESS, G., AND M. RABIN (2002): “Understanding Social Preferences with Simple Tests,” *Quarterly Journal of Economics*, 117(3), 817–870.
- CHOU, E., M. MCCONNELL, R. NAGEL, AND C. PLOTT (2009): “The control of game form recognition in experiments: understanding dominant strategy failures in a simple two person guessing game,” *Experimental Economics*, 12(2), 159–179.

- CIALDINI, C., J. BASSETT, R. MILLER, AND J. MILLER (1978): "Low-ball procedure for producing compliance: Commitment, then cost," *Journal of Personality and Social Psychology*, 36(5), 463–476.
- CIALDINI, R., AND B. SAGARIN (2005): "Interpersonal influence," in *Persuasion: Psychological insights and perspectives*, ed. by T. Brock, and M. Green, pp. 143–169. Sage Press, Newbury Park, CA.
- CLARK, K., S. KAY, AND M. SEFTON (2001): "When are Nash equilibria self-enforcing? An experimental analysis," *International Journal of Game Theory*, 29(4), 495–515.
- COOPER, D. J., AND J. B. VAN HUYCK (2003): "Evidence on the equivalence of the strategic and extensive form representation of games," *Journal of Economic Theory*, 110(2), 290–308.
- COOPER, R., D. V. DEJONG, R. FORSYTHE, AND T. W. ROSS (1989): "Communication in the Battle of the Sexes Game: Some Experimental Results," *The RAND Journal of Economics*, 20(4), pp. 568–587.
- (1990): "Selection Criteria in Coordination Games: Some Experimental Results," *American Economic Review*, 80(1), 218–33.
- (1992): "Communication in Coordination Games," *The Quarterly Journal of Economics*, 107(2), 739–71.
- CRAWFORD, V. (1998): "A Survey of Experiments on Communication via Cheap Talk," *Journal of Economic Theory*, 78(2), 286–298.
- CRAWFORD, V. P., AND J. SOBEL (1982): "Strategic Information Transmission," *Econometrica*, 50(6), 1431–51.
- CZAP, H. J., N. V. CZAP, M. KHACHATURYAN, M. E. BURBACH, AND G. D. LYNNE (2011): "Smiley or Frowney: The effect of emotions and framing in a downstream water pollution game," Discussion paper.
- DEVETAG, G., AND A. ORTMANN (2007): "When and why? A critical survey on coordination failure in the laboratory," *Experimental Economics*, 10(3), 331–344.
- DUFFY, J., AND N. FELTOVICH (2002): "Do Actions Speak Louder Than Words? An Experimental Comparison of Observation and Cheap Talk," *Games and Economic Behavior*, 39(1), 1–27.
- (2006): "Words, Deeds, and Lies: Strategic Behaviour in Games with Multiple Signals," *Review of Economic Studies*, 73(3), 669–688.
- DUGAR, S. (2010): "Nonmonetary sanctions and rewards in an experimental coordination game," *Journal of Economic Behavior & Organization*, 73(3), 377–386.
- (2013): "Non-monetary incentives and opportunistic behavior: evidence from a laboratory public good game," *Economic Inquiry*, 51(2), 1374–1388.
- ELLINGSEN, T., AND M. JOHANNESSON (2004): "Promises, Threats and Fairness," *Economic Journal*, 114(495), 397–420.



- (2008): “Anticipated verbal feedback induces altruistic behavior,” *Evolution and Human Behavior*, 29(2), 100–105.
- ELLINGSEN, T., AND R. ÖSTLING (2010): “When Does Communication Improve Coordination?,” *American Economic Review*, 100(4), 1695–1724.
- EMBREY, M. S., G. R. FRÉCHETTE, AND S. F. LEHRER (2012): “Bargaining and Reputation: An Experiments on Bargaining in the Presence of Behavioral Types,” mimeo, New York University.
- ENGELMANN, D., AND M. STROBEL (2004): “Inequality Aversion, Efficiency, and Maximin Preferences in Simple Distribution Experiments,” *American Economic Review*, 94(4), 857–869.
- FALK, A., AND J. J. HECKMAN (2009): “Lab Experiments Are a Major Source of Knowledge in the Social Sciences,” *Science*, 326(5952), 535–538.
- FARRELL, J. (1987): “Cheap Talk, Coordination, and Entry,” *RAND Journal of Economics*, 18(1), 34–39.
- (1988): “Communication, coordination and Nash equilibrium,” *Economics Letters*, 27(3), 209–214.
- FARRELL, J., AND M. RABIN (1996): “Cheap Talk,” *Journal of Economic Perspectives*, 10(3), 103–118.
- FEHR, E., M. NAEF, AND K. M. SCHMIDT (2006): “Inequality Aversion, Efficiency, and Maximin Preferences in Simple Distribution Experiments: Comment,” *American Economic Review*, 96(5), 1912–1917.
- FEHR, E., AND K. M. SCHMIDT (1999): “A Theory of Fairness, Competition, and Cooperation,” *Quarterly Journal of Economics*, 114(3), 817–868.
- FELTOVICH, N., AND J. SWIERZBINSKI (2011): “The role of strategic uncertainty in games: An experimental study of cheap talk and contracts in the Nash demand game,” *European Economic Review*, 55(4), 554–574.
- FERRARO, P. J., AND C. A. VOSSLER (2010): “The Source and Significance of Confusion in Public Goods Experiments,” *The B.E. Journal of Economic Analysis & Policy*, 10(1), 53.
- FISCHBACHER, U., S. GACHTER, AND E. FEHR (2001): “Are people conditionally cooperative? Evidence from a public goods experiment,” *Economics Letters*, 71(3), 397–404.
- GOEREE, J. K., AND C. A. HOLT (2001): “Ten Little Treasures of Game Theory and Ten Intuitive Contradictions,” *American Economic Review*, 91(5), 1402–1422.
- GRANDJEAN, I., AND N. GUÉGUEN (2011): “Testing a binding communication strategy in a company: How could persuasive information be more efficient,” *Social Behavior and Personality*, 39(9), 1209–1216.
- GREINER, B. (2004): “An Online Recruitment System for Economic Experiments,” in *Forschung und wissenschaftliches Rechnen 2003. GWDG Bericht 63*, ed. by K. Kremer, and V. Macho, pp. 79–93. Ges. für Wiss. Datenverarbeitung, Göttingen.

- GUÉGUEN, N., R.-V. JOULE, S. HALIMI-FALKOWICZ, A. PASCUAL, J. FISCHER-LOKOU, AND M. DUFOURCQ-BRANA (2011): “I’m Free but I’ll Comply With Your Request : Generalization and Multidimensional Effects of “Evoking freedom” Technique,” *Journal of Applied Social Psychology*, Forthcoming.
- HANAKI, N., N. JACQUEMET, S. LUCHINI, AND A. ZYLBERSZTEJN (2013): “Bounded Rationality and Strategic Uncertainty in a Simple Dominance Solvable Game,” Discussion paper.
- HARDIN, G. (1968): “The Tragedy of the Commons,” *Science*, 162, 1243–1248.
- HARRISON, G. W., AND J. HIRSHLEIFER (1989): “An Experimental Evaluation of Weakest Link/Best Shot Models of Public Goods,” *Journal of Political Economy*, 97(1), 201–25.
- HARSANYI, J. C., AND R. SELTEN (1988): *A General Theory of Equilibrium Selection in Games*, vol. 1 of *MIT Press Books*. The MIT Press.
- HAYEK, F. A. (1945): “The Use of Knowledge in Society,” *The American Economic Review*, 35(4), pp. 519–530.
- HENRICH, J., S. J. HEINE, AND A. NORENZAYAN (2010): “The Weirdest People in the World?,” *Behavioral and Brain Sciences*, 33(2/3), 61–83.
- HOLLÄNDER, H. (1990): “A Social Exchange Approach to Voluntary Cooperation,” *The American Economic Review*, 80(5), pp. 1157–1167.
- ISAAC, R. M., AND J. M. WALKER (1988): “Communication and Free-Riding Behavior: The Voluntary Contribution Mechanism,” *Economic Inquiry*, 26(4), 585–608.
- IVANOV, A., D. LEVIN, AND M. NIEDERLE (2010): “Can relaxation of beliefs rationalize the winner’s curse?: An experimental study,” *Econometrica*, 78(4), 1435–1452.
- JACQUEMET, N., A. JAMES, S. LUCHINI, AND J. SHOGREN (2010): “Referenda under Oath,” Université Paris1 Panthéon-Sorbonne (Post-Print and Working Papers) halshs-00490448, HAL.
- JACQUEMET, N., R.-V. JOULE, S. LUCHINI, AND J. SHOGREN (2009): “Earned Wealth, Engaged Bidders? Evidence from a second price auction,” *Economics Letters*, 105(1), 36–38.
- (2012): “Eliciting Preferences under Oath,” *Journal of Environmental Economics and Management*, Forthcoming.
- JACQUEMET, N., R.-V. JOULE, S. LUCHINI, AND J. F. SHOGREN (2013): “Preference elicitation under oath,” *Journal of Environmental Economics and Management*, 65(1), 110 – 132.
- JACQUEMET, N., S. LUCHINI, J. SHOGREN, AND A. ZYLBERSZTEJN (2011): “Coordination with Communication under Oath,” *GREQAM WP*, 2011(49).
- JACQUEMET, N., AND A. ZYLBERSZTEJN (2011): “What drives failure to maximize payoffs in the lab? A test of the inequality aversion hypothesis?,” *CES Working Paper*, 2011(36).

- (2012): “Learning, words and actions: experimental evidence on coordination-improving information,” *B.E. Journal of Theoretical Economics*, Forthcoming.
- JOULE, R., AND J. BEAUVOIS (1998): *La soumission librement consentie*. Presses Universitaires de France, Paris.
- JOULE, R.-V., F. GIRANDOLA, AND F. BERNARD (2007): “How Can People Be Induced to Willingly Change Their Behavior? The Path from Persuasive Communication to Binding Communication,” *Social and Personality Psychology Compass*, 1(1), 493–505.
- KAMECKE, U. (1997): “note: Rotations: Matching Schemes that Efficiently Preserve the Best Reply Structure of a One Shot Game,” *International Journal of Game Theory*, 26(3), 409–417.
- KATZEV, R., AND T. WANG (1994): “Can commitment change behavior ? A case study of environmental actions,” *Journal of Social Behavior and Personality*, 9, 13–26.
- KIESLER, C. (1971): *The psychology of commitment. Experiments linking behavior to belief*. Academic Press, New York.
- KIESLER, C., AND J. SAKUMURA (1966): “A test of a model for commitment,” *Journal of Personality and Social Psychology*, 3(3), 349–353.
- LÓPEZ-PÉREZ, R., AND M. VORSATZ (2010): “On approval and disapproval: Theory and experiments,” *Journal of Economic Psychology*, 31(4), 527–541.
- LUCE, R. D., AND H. RAIFFA (1957): *Games and decisions : introduction and critical survey / R. Duncan Luce and Howard Raiffa*. Wiley New York.
- MASCLET, D., C. NOUSSAIR, S. TUCKER, AND M.-C. VILLEVAL (2003): “Monetary and Nonmonetary Punishment in the Voluntary Contributions Mechanism,” *American Economic Review*, 93(1), 366–380.
- MASCLET, D., C. NOUSSAIR, AND M.-C. VILLEVAL (forthcoming): “Threat and punishment in public good experiment,” *Economic Inquiry*.
- OSTROM, E. (1990): *Governing the commons-The evolution of institutions for collective actions*. Political economy of institutions and decisions.
- PALLACK, M., D. COOK, AND J. SULLIVAN (1980): “Commitment and energy conservation,” in *Applied Social Psychology Annual*, ed. by L. Bickman, pp. 235–253. Beverly Hills, CA: Sage.
- PARKHURST, G. M., J. F. SHOGREN, AND C. BASTIAN (2004): “Repetition, Communication, and Coordination Failure,” *Experimental Economics*, 7(2), 141–152.
- PEETERS, R., AND M. VORSATZ (2009): “Immaterial rewards and sanctions in a voluntary contribution experiment,” Discussion paper.
- (forthcoming): “Immaterial rewards and sanctions in a voluntary contribution experiment,” *Economic Inquiry*.

- PLOTT, C., AND K. ZEILER (2005): "The Willingness to Pay-Willingness to Accept Gap, the "Endowment Effect," Subject Misconceptions, and Experimental Procedures for Eliciting Valuations," *American Economic Review*, 85, 530–45.
- REGE, M., AND K. TELLE (2004): "The impact of social approval and framing on cooperation in public good situations," *Journal of Public Economics*, 88(7-8), 1625–1644.
- RIEDL, A., I. ROHDE, AND M. STROBEL (2011): "Efficient Coordination in Weakest-Link Games," *CESifo Working Paper Series*, 3685, Available at SSRN: <http://ssrn.com/abstract=1980063>.
- ROSENTHAL, R. W. (1981): "Games of perfect information, predatory pricing and the chain-store paradox," *Journal of Economic Theory*, 25(1), 92–100.
- SALLY, D. (1995): "Conversation and Cooperation in Social Dilemmas: A Meta-Analysis of Experiments from 1958 to 1992," *Rationality and Society*, 7(1), 58–92.
- SAMUELSON, P. A. (1954): "The Pure Theory of Public Expenditure," *The Review of Economics and Statistics*, 36(4), pp. 387–389.
- SCHELLING, T. (1960): *The strategy of conflict*. Harvard University Press, Cambridge, MA.
- SEKHON, J. (2011): "Multivariate and Propensity Score Matching Software with Automated Balance Optimization," *Journal of Statistical Software*, 42(7), 1–52.
- SELTEN, R. (1975): "Reexamination of the perfectness concept for equilibrium points in extensive games," *International Journal of Game Theory*, 4(1), 25–55.
- TRICOMI, E., A. RANGEL, C. F. CAMERER, AND J. P. O'DOHERTY (2010): "Neural evidence for inequality-averse social preferences," *Nature*, 463, 1089–1091.
- VANBERG, C. (2008): "Why Do People Keep Their Promises? An Experimental Test of Two Explanations -super-1," *Econometrica*, 76(6), 1467–1480.
- WANG, T., AND R. KATSEV (1990): "Groupe commitment and resource conservation : two field experiments on promoting recycling," *Journal of Applied Social Psychology*, 20, 265–275.
- WILLIAMS, R. (2000): "A Note on Robust Variance Estimation for Cluster-Correlated Data," *Biometrics*, 56(2), 645–646.
- WILSON, R. K., AND J. SELL (1997): "'Liar, Liar...': Cheap Talk and Reputation in Repeated Public Goods Settings," *The Journal of Conflict Resolution*, 41(5), pp. 695–717.
- WOOLDRIDGE, J. M. (2003): "Cluster-Sample Methods in Applied Econometrics," *American Economic Review*, 93(2), 133–138.
- XIAO, E., AND D. HOUSER (2009): "Avoiding the sharp tongue: Anticipated written messages promote fair economic exchange," *Journal of Economic Psychology*, 30(3), 393–404.

- ZEILIGER, R. (2000): “A presentation of Regate, Internet based Software for Experimental Economics,”  
*<http://www.gate.cnrs.fr/zeiliger/regate/RegateIntro.ppt>, GATE.*
- ZIZZO, D. J., AND A. J. OSWALD (2001): “Are People Willing to Pay to Reduce Others’ Incomes?,”  
*Annales d’Economie et de Statistique*, (63-64), 39–65.